The Journal of
Applied Psychology

Edited by

JAMES P. PORTER
Ohio University

WILLIAM F. BOOK
Indiana University

With a Board of Co-operating Editors

VOLUME IX
1925

PUBLISHED BY THE EDITORS

Reprinted with the permission of the original publishers.
JOHNSON REPRINT CORPORATION
KRAUS REPRINT CORPORATION

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. V. Bingham, Carnegie Institute of Technology; June E. Downey, University of Wyoming; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; Joseph Jastrow, University of Wisconsin; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College, Henry C. McComas, Princeton University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, Harvard Graduate School of Business Administration; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; J. E. Wallace Wallin, Teachers College, Miami University; G. M. Whipple, University of Michigan; Helen Thompson Woolley, Merrill-Palmer School, Detroit.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeal. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or musical instruments, sports, etc.

Copyright, 1926, by James P. Porter

CONTENTS

IRVING ALLEN: *See* Kitson and Allen.
ADA HART ARLITT: *See* Dowd and Arlitt.
GRACE ARTHUR:
A New Point Performance Scale 390
MARION A. BILLS:
Social Status of the Clerical Worker and His Permanence on the Job 424
GEORGE C. BRANDENBURG:
Analyzing Personality, Part I 139
Personality and Vocational Achievement, Part II 281
H. E. BURTT AND E. M. DOBELL:
The Curve of Forgetting for Advertising Material 5
E. M. DOBELL: *See* Burtt and Dobell.
CONSTANCE E. DOWD AND ADA HART ARLITT:
The Relative Transfer Effects of Supervised Play and Formal Gymnastics 215
MAX FREYD:
The Statistical Viewpoint in Vocational Selection 349
DOUGLAS FRYER:
Types of Work. A Note on Vocational Interest 304
THOMAS R. GARTH:
The Intelligence of Full-Blood Indians 382
J. P. GUILFORD: *See* Hyde and Guilford.
RALPH K. HOITSMA:
The Reliability and Relationships of the Colgate Mental Hygiene Test 293
H. L. HOLLINGWORTH:
When is a Man Intoxicated? 122
JOSEPH L. HOLMES:
The Free Association Method as a Measure of the Efficiency of Advertising 60
WINIFRED F. HYDE AND J. P. GUILFORD:
A Test for Classification of Students in Chemistry 196
HARRY D. KITSON AND IRVING ALLEN:
Minor Studies in the Psychology of Advertising from the Laboratory of Indiana University 307
T. A. LANGLIE: *See* Paterson and Langlie.

Gladys M. Lowe and Myra E. Shimberg:
A Critique of the Fables as a Moral Judgment Test.............. 53
Grace E. Manson:
Differences in Intelligence Tests. The Relative Difficulty of Types of Questions.. 156
Personality Differences in Intelligence Test Performance....... 230
Mark A. May:
The Present Status of the Will-Temperament Tests........... 29
J. B. Miner:
The Prediction of a Disparity Between Scholarship and Intelligence.. 357
Katharine Murdoch:
A Study of the Comparative Value of Nine Performance Tests. . 364
H. K. Nixon:
Two Studies of Attention to Advertisements.................. 176
Frances E. Otis:
Phenomenal Memory in Its Bearing upon Various Mental Tests. 311
Donald G. Paterson and T. A. Langlie:
Empirical Data on the Scoring of True-False Tests............. 339
A. T. Poffenberger:
School Achievement and Success in Life...................... 22
H. H. Remmers and L. A. Thompson:
A Note on Motor Activity as Conditioned by Emotional States. 417
W. A. Rorison: *See* Warden, Rorison and Solinger.
Sadie Myers Shellow:
A Suggestion for an Experiment on the Effect of Note-Taking in Lecture Courses.. 60
Eugene Shen:
The Influence of Friendship upon Personal Ratings............ 66
Myra E. Shimberg: *See* Lowe and Shimberg.
S. B. Solinger: *See* Warden, Rorison and Solinger.
Sara M. Stinchfield:
The Speech of Five Hundred Freshmen College Women........ 109
Edward K. Strong, Jr.:
Theories of Selling.. 75
Dagny Sunne:
Personality Tests: White and Negro Adolescents.............. 256
Percival M. Symonds:
Notes on Rating... 188
L. A. Thompson, Jr.: *See* Remmers and Thompson.
E. L. Thorndike:
On the Provision of Alternative Forms of Examinations Equal in Difficulty.. 1

MORRIS S. VITELES:
The Clinical Viewpoint in Vocational Selection................ 131
C. J. WARDEN, W. A. ROBISON AND S. B. SOLINGER:
Tendencies in the Use of Headlines in Magazine Advertising.... 225
HERMAN H. YOUNG:
A Speech Clinic Case with Misconduct as a By-Product........ 371

ON THE PROVISION OF ALTERNATIVE FORMS OF EXAMINATIONS EQUAL IN DIFFICULTY

E. L. THORNDIKE

Teachers College, Columbia University

It is obviously important that alternative forms of psychological and educational tests or examinations be of approximately the same difficulty and that the amounts of what differences they do have be known. For example, the enormous variation in difficulty among the examinations in the same subject used by the College Entrance Examination Board, together with the fact that the amounts of these differences are known too late to be corrected for, has been their most notorious weakness. It is now proposed to report the percentile position of each candidate as well as his actual score. This will be a considerable alleviation, but new dangers are introduced by the substitution of relative positions for measures of actual ability and achievement. As soon as it becomes known that a college will accept a certain top fraction of candidates who take a College Entrance Board Examination, or that a civil service list will be filled by a certain top fraction of examinees, there will be a strong temptation to increase the number of candidates. Also, the measurement of any change in the general quality from year to year is impossible by percentiles computed separately for each examination. It is consequently desirable to know what may be expected in the way of similarity in active examination-scores from a reasonable amount of care and effort in preparing and equating the alternative forms. The facts concerning fifteen alternative forms of a certain intelligence examination were reported in this JOURNAL (Vol. IV, pp. 283–288). The present note concerns twelve alternative forms of the Thorndike Intelligence Examination for High School Graduates.

Using the same methods of determining difficulty as those described in the earlier report (chiefly the taking of alternative *forms* by *random* halves of a group), it appears that the difficulty of any form (for the median candidate for entrance to our best colleges) is rather close to that of any other. To bring the score in any one of the seven examinations of Series 1919–1924 to equality with the score in any other, we have to add as follows:

To the score in Examination of September 1919, DEBB, add -3
To the score in Examination of September 1920, IMDD, add $+3$
To the score in Examination of September 1921, LOEE, add -1
To the score in Examination of September 1922, DNFF, add 0
To the score in Examination of September 1923, EJC(AA) add $+1$
To the score in Examination of September 1924, BCGC, add $+\frac{1}{2}$
To the score in Examination (extra) GHAH, add $-\frac{1}{2}$

The average score for candidates is about 75, and the mean square deviation is about 14. The average difference of one form of the examination from another in difficulty (2.2) is thus about $\frac{1}{6}$ of the mean square deviation of the group. The bulk of the differences is due to the first two examinations. In the last five, after time had been available for experimentation, the differences are greatly reduced.[1] The five examinations prepared for 1925–1930 differ one from another in difficulty, by only a quarter of one score point on the average.

This close equivalence was obtained only after much experimentation. In the 1919–1924 series, where the examination *consists of twenty-one different sorts of tasks, with from five* to sixty elementary tasks of each sort, the experimentation required was simply to measure the comparative difficulty of the different forms of Part I, Part II, and Part III, and put them together in combinations such that the chance greater ease of one would be balanced by the chance greater difficulty of another. With the 1925–1930 series, where there are only fourteen sorts of tasks, the experimentation consisted in much preliminary equalizations of tests and parts as well as in final arrangement of Parts in combinations. Also certain "Parts" of extreme ease or difficulty were discarded, the fifteen parts used being a selection from seventeen. In the preliminary work, many forms of the reading tests were discarded. The selection of the forty tests used in the five alternative forms of Part III, for example, was from over fifty which were tried out, each with from fifteen to a hundred individuals, before any form of Part III was constructed.

---

[1] With earlier experimentation, II-B and II-D could have been interchanged, making the 1920 and 1921 examinations almost exactly alike and equal to the average in difficulty.

The amount of time spent in equating the five forms of the 1925–1930 series, after the test materials had all been selected and made into provisional series by random division, was at least eight thousand hours of time spent by subjects in taking tests, at least four thousand hours spent in scoring and tabulating scores. This time could have been reduced by four-fifths, if a large number of individuals of the sort desired could have been induced to take all the different alternative forms, doing their very best with each. Our efforts in this direction, however, met with very little success. Very few college freshmen were willing to take such a series, though the payment offered was double that given for student clerical work. Those who did take the tests did not do so in a way to assure one that they were trying to do their best. It seemed wisest therefore to use groups where the examination was *bona fide*, giving two Parts or two tests whose difficulty was to be equated to random halves of the group.

Doubtless with more experience and wiser planning, the equating could have been more economically done, but the task of equating satisfactorily two-hour or three-hour examinations in Plane Geometry or French, or fitness to be a general clerk or a rural delivery postman, will require much labor.

There is danger that when examining boards appreciate the amount of labor required to make a series of examinations equal in difficulty, they will sacrifice the quality of the examinations in order to reduce this labor. The procedure should, of course, be to choose all the materials for a series of examinations for their *significance*, that is, the *validity* correlation of their composite with the criterion. The maximum convenience in giving, scoring, and arranging into alternative forms should then be secured, but at no cost to the significance of the scores.

THE CURVE OF FORGETTING FOR ADVERTISING MATERIAL

H. E. BURTT AND E. M. DOBELL

Ohio State University

THE PROBLEM

The curve of forgetting was determined originally by Ebbinghaus.[1] He studied nonsense syllables until he could repeat them correctly and then measured the rate of forgetting by the amount of time required for relearning them at different intervals after the original memorizing. Somewhat similar studies with nonsense syllables, poetry, or letters have been made by Radossawljewitsch and by Bean.[2] The results are usually shown by plotting time on the abscissa and retention on the ordinate. Such curves in the foregoing studies all agree in showing a rather abrupt drop at the outset followed by a much more gradual decline. For instance with nonsense syllables as much was forgotten in the first twenty minutes as in the following thirty days, and with poetry as much was lost in two days as in the next twenty-five. This suggests the desirability, if material is to be presented repeatedly, of having the earlier repetitions closer together than the later in order to strengthen the associations before they have become greatly weakened.

It seems plausible that these same principles would apply to memory for advertising material. It has been suggested that, if an advertisement is to be repeated or some type of follow-up procedure used, the early repetitions should come closer together than the later. For instance Kitson[3] suggests a distribution of eight advertisements through a year's campaign with the early ones a month apart and the later ones two months. Similarly Hollingworth[4] proposes a series of appeals directed to the same person at intervals of 2 days, 5 days, 10 days and 20 days.

---

[1] Ebbinghaus, H. Ueber das Gedachtniss. 1885.

[2] Radossawljewitsch, P. R. Das Behalten und Vergessen bei Kindern fund Erwachsenen nach experimentallen Untersuchungen. 1907.

Bean, C. H. The Curve of Forgetting. *Archives of Psychology*, no. 21. For a brief summary of this work and a chart showing typical curves of forgetting cf. Starch, D. Educational Psychology, 1920, p. 156.

While this generalization from laboratory experiments with nonsense syllables, words or poetry to advertising procedure involving concrete names of commodities or brands is probably warranted it seemed well to test it experimentally. The present study accordingly, required the subject to associate names of commodities with fictitious trade names of those commodities and determined the curve of forgetting. In some series follow-up procedure—i.e., a second repetition of the names—was introduced at different intervals after the original presentation in order to discover the relative merits of early and late follow-up.

SERIES I

Material

The material for the first series consisted of 100 pairs of words. The first word of each pair was the name of a commodity and the second a fictitious brand or trade name of that commodity, such as:

Perfume.............Pettal, China............Mangone

Fictitious trade names were used to avoid error due to certain commercial names being more familiar to some subjects than to others. These 100 pairs were selected as a result of preliminary experiments with a few subjects to discard any that involved unusual or catchy associations. They were typed in random order on a wide adding machine ribbon with the pairs 1 inch apart.

---

[3] Kitson, H. D. The Mind of the Buyer. 1921, p. 48.

[4] Hollingworth, H. L. Advertising and Selling. 1920, p. 205.

A Balopticon was adapted for the presentation of the words. The ribbon was fed across the illuminated field between guides. This field was covered by a piece of metal containing a slot about 20 by 70 mm. through which the pairs of words were visible as they passed underneath and were thus projected on the screen.

Mimeographed blanks were provided containing selected lists of the original commodities such as:

Perfume............................
China...............................

Other blanks had the name of each commodity followed by four alternative trade names, one of which had occurred with it in the original list, such as:

Perfume......................Arden; Pettal; Lotus; Blossom
China.........................Gosser; Pope; Curtis; Mangone

Method

The group method of experiment was used. The subjects sat in a dark room and observed the material as it was projected on the screen by means of the Balopticon. The ribbon containing the pairs of words was pulled along between the guides by the experimenter. A metronome somewhat muffled indicated to him the time and every 4 seconds he pulled the ribbon about 1 inch thus exposing a new pair of words. This motion had been practised previously and with the exposure of 4 seconds probably did not involve a significant error. The list was shown a second time in identical manner immediately after the first presentation. It seemed better to do this than to give one presentation with 8 second exposure because of possible lapses of attention.

The subjects were instructed as follows:

> This is an experiment upon memory for trade names. You will be shown in succession pairs of names like "soap—Ivory," "automobile—Ford." The first name in each pair is the name of a commodity and the second name in each pair is that of a brand. Your are to try to remem-

ber the names of each pair together so that subsequently when you see the name of the commodity you will be able to recall the corresponding brand. For instance, you should remember "soap" and "ivory," together so that when later shown "soap" you will recall "ivory." The brand names in the actual experiment, however, are mostly unfamiliar. There are 100 pairs in the list and you will see each for 4 seconds. You will not be expected to remember every one but to get all you can.

Immediately after the words had been presented twice the lights were turned on and blanks distributed for the initial recall. Each subject received a blank containing 20 names of commodities that had appeared in the screen. However, all the subjects did not receive the same list. The 100 words were divided at random into 5 groups of 20 and each group was mimeographed on a separate blank. Moreover the subjects were divided into 5 squads of approximately equal size. In the initial test squad A was given list 1, squad B list 2, squad C list 3, squad D list 4, and squad E list 5. When another test was given later (infra) squad A was given list 2, squad B list 3, etc. Likewise for still later tests each squad had a still different list. Thus each squad recalled in subsequent tests words upon which it had not been tested previously. In each test, however, every one of the 100 words was given to some squad. In this way average results for all subjects on different tests tended to eliminate any errors due to difficulty of the words.

In the initial test each squad was first given the list of 20 names of commodities and instructed to write after each the name of the brand that had been presented with it on the screen. No time limit was set for this recall but all the subjects finished in less than 4 minutes. As soon as the blanks were collected a second set was distributed. Each squad now received the same names of commodities as before but each was followed by four alternative brands and the subjects checked the alternative that had been presented with the commodity originally. No time limit was set but all finished within 3 minutes. In subsequent tests the subjects worked more rapidly so that 3 minutes was sufficient both for recall and for recognition. Thus on the first day there was available for each subject his score in recalling and in recognizing 20 trade names and results were obtained for 5 squads each with a different group of 20 names.

The purpose of this series was to determine the curve of forgetting without any follow-up procedure. One week after the first test, each squad was given blanks containing 20 of the original names of commodities, but a different 20 from those recalled by that squad on the first day. The original list was *not* presented on the screen at this time nor subsequently during this series. Each squad recalled the associated trade name as in the first test. The blanks were collected and the corresponding ones distributed for the recognition test in exactly the same manner as before. A week later each squad was tested on a third list of 20 words in identical fashion. Similarly after 3 and 4 weeks other tests were made.

The subjects were 58 students in an advanced class taught by the senior writer who likewise conducted the actual experiment. In this and subsequent series the results of subjects who did not participate in the entire experiment were discarded.

Results

The subjects' blanks were scored number correct. There was of course a frequent tendency to recall a trade name that had been presented with some other commodity than the one in question but no consideration was given to such cases. Likewise mere similarity in meaning or in some of the letters was insufficient to give a correct score unless there was absolute identity.

The results are summarized in table 1. Each figure indicates the per cent of the list of 20 recalled or recognized at the time specified. The successive rows indicate the average[5] results for the different squads and the average of all squads. The columns indicate results for the initial test, for the test after 1 week, 2 weeks, etc. The left portion of the table gives results for recall, i.e., where the subject wrote after the stimulus word the one that had been shown with it previously. The right portion of the table gives results for recognition, i.e., where the subject selected one of the four alternatives. For instance, immediately after the initial presentation squad A on the average recalled 40 per cent of the words on which it was tested while it recognized 74 per cent. A week later this squad on the average recalled only 5 per cent of the words on which it was tested but recognized 71 per cent correctly. Two weeks after the initial presentation this squad was able to recall only 1 per cent but recognized 66 per cent. The column at the extreme right gives the number of subjects in each squad. The bottom row gives the averages (unweighted) of the figures in the column above it. The same results are shown graphically in figure 1. The abscissa represents time, i.e., initial, 1 week, 2 weeks, etc. The ordinate represents per cent of retention—recall in the lower set of curves and recognition in the upper. A separate curve is made for each squad and for the average of all squads as indicated in the key.

---

[5] The arithmetic mean is the average used throughout.

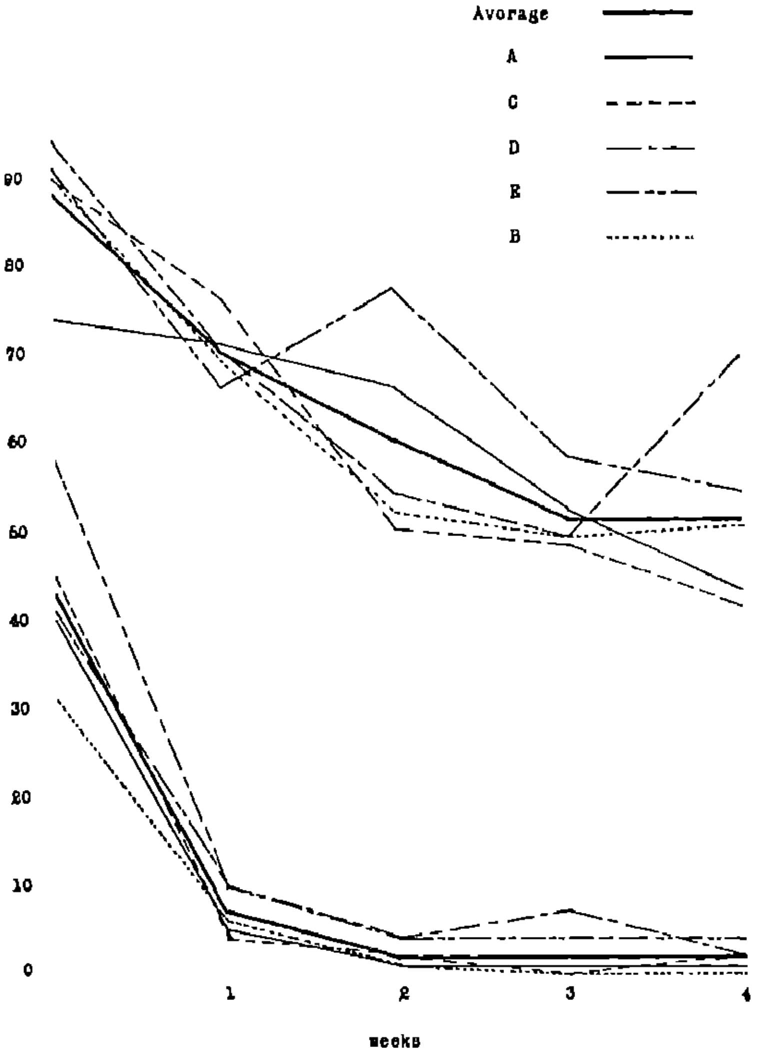
TABLE 1

Per cent of retention after different intervals

SQUAD	RECALL					RECOGNITION					N
	Initial	1 week	2 weeks	3 weeks	4 weeks	Initial	1 week	2 weeks	3 weeks	4 weeks	
A	40	5	1	1	1	74	71	66	52	43	10
B	31	6	1	0	0	90	69	52	49	50	14
C	45	4	2	0	2	90	76	50	48	41	12
D	58	10	4	7	2	94	70	54	40	69	9
E	41	10	4	4	4	91	60	77	68	54	13
Average........	43	7	2	2	2	88	70	60	51	51	

The results show the typical progress of forgetting. The average retention drops from 43 to 7 per cent within one week and within two weeks drop to 2 per cent. In the recognition method the drop is not as abrupt but still very marked. The

Figure I



curves appear similar to those found in previous studies of forgetting with other kinds of material.

This preliminary series then confirms the usual assumption that advertising material follows the same principles as does other memory material with respect to the curve of forgetting. This is especially true in recall and this is perhaps the aspect that is most important for the advertiser. He is frequently attempting to have a commodity associated with his trade name so strongly that when subsequently the need for that commodity arises the prospect will think of the proper trade name. This sort of association in our results weakens very rapidly with time.

SERIES II

Material

The material for this series was the same as that in Series I, viz., the list of 100 pairs of names typed on a ribbon. The blanks for testing recall and recognition however were made in units of 10 rather than 20 words. The makeup of these lists was not entirely random. Three of them involved the first 30 words of the original list arranged at random; three more involved the next 30 words of the original list likewise in random order and the remaining 40 of the original list were taken at random to make the other 4 lists. This arrangement was necessary in order to allow a consecutive portion of the original list to be subsequently presented again (cf. infra).

Method

The method of initial presentation and the instructions were identical with Series I. The mimeographed blanks for recall and recognition were likewise used in identical manner. However this series studied the effect of "follow up," i.e., a second presentation of the material at a subsequent time and particularly the effect of having this follow-up early or late in the series. The first two tests were also closer together than in Series I inasmuch as it was suspected that the curve would approach its minimum in less than a week. The subjects were divided into 3 squads.

The entire list of 100 was presented twice in immediate succession at the rate of 4 seconds per word on February 5. Each squad was tested by recall and recognition methods on two lists of 10 words each. On February 8, 30 pairs of the original list (numbers 1 to 30) were presented once at the usual rate. Immediately each squad was tested on 10 words from this 30 and also on 10 words of the original list that had not been presented since February 5 and that had not been previously tested with the squad in question. On February 15, 30 more pairs of the original list (numbers 31 to 60) were presented once in the usual way. Immediately each squad was tested on 10 of these, on 10 that had been presented only on the 5th but never tested until the 15th with the squad in question and also on 10 that had been presented on the 5th and again on the 8th but not tested till the 15th with that squad. On the 21st no further presentation was given but each squad was tested on 30 words with which it had not been tested previously—of which 10 had been given on the 5th, 10 on the 5th and 8th and 10 on the 5th and 10th. A careful grouping of the words and arrangement of schedule made it possible to carry through this program with no two squads having the same list under the same conditions and with each list occurring under every condition for some squad. It was possible to obtain from this data three curves for each squad—the ordinary curve of forgetting as in Series I, a curve of forgetting with early follow-up—i.e., on the 8th and a curve of forgetting with late follow-up, i.e., on the 15th.

As before the lists were scored on the basis of number absolutely correct. The subjects were 41 students in an advanced class taught by the senior writer who also conducted the experiment.

Results

The results are summarized in table 2. The figures represent the per cent of the list recalled or recognized under the given conditions. For each squad the first row indicates results for words presented only at the outset, i.e., the ordinary curve of forgetting as in Series I. The second row gives memory for words presented initially and also after 3 days and the third row for words presented initially and after 10 days. The columns indicate the day on which the words were tested. By way of illustration the recall results for squad A will be explained. With words of the original list that were shown only on the first day 49 per cent were recalled immediately. When tested on others of these words after 3 days only 9 per cent were recalled; after 10 days only 1 per cent and after 16 days 2 per cent. This is the typical curve of forgetting. With words that were shown at the outset and a second time after 3 days and tested immediately thereafter 58 per cent were recalled, while other words presented similarly but not tested until 10 days after the beginning had 23 per cent retention and still others similarly presented but not tested until the last

TABLE 2

Per cent of retention after different intervals with early and late follow up

SQUAD		RECALL				RECOGNITION				N
	Presentation	0 days	3 days	10 days	16 days	0 days	3 days	10 days	16 days	
A	Initial	49	9	1	2	88	71	54	44	14
	Initial and 3 days		58	23	10		98	93	70	
	Initial and 10 days			42	13			93	78	
B	Initial	47	16	8	6	96	82	66	63	12
	Initial and 3 days		55	18	15		100	88	79	
	Initial and 10 days			40	10			88	81	
C	Initial	59	9	9	8	94	73	73	84	15
	Initial and 3 days		57	25	22		97	89	91	
	Initial and 10 days			54	21			96	84	
Average	Initial	52	11	6	5	93	75	64	64	
	Initial and 3 days		57	22	16		99	90	80	
	Initial and 10 days			48	15			92	81	

day of the series had a 10 per cent retention. With words shown at the outset and a second time after 10 days and tested immediately thereafter 42 per cent were recalled while other words presented similarly but not tested till 16 days after the beginning had a retention of 13 per cent. Similar data is given for recognition and for all three squads. The bottom block gives the averages (unweighted) of the corresponding figures above. The extreme right column gives the numbers in each squad.

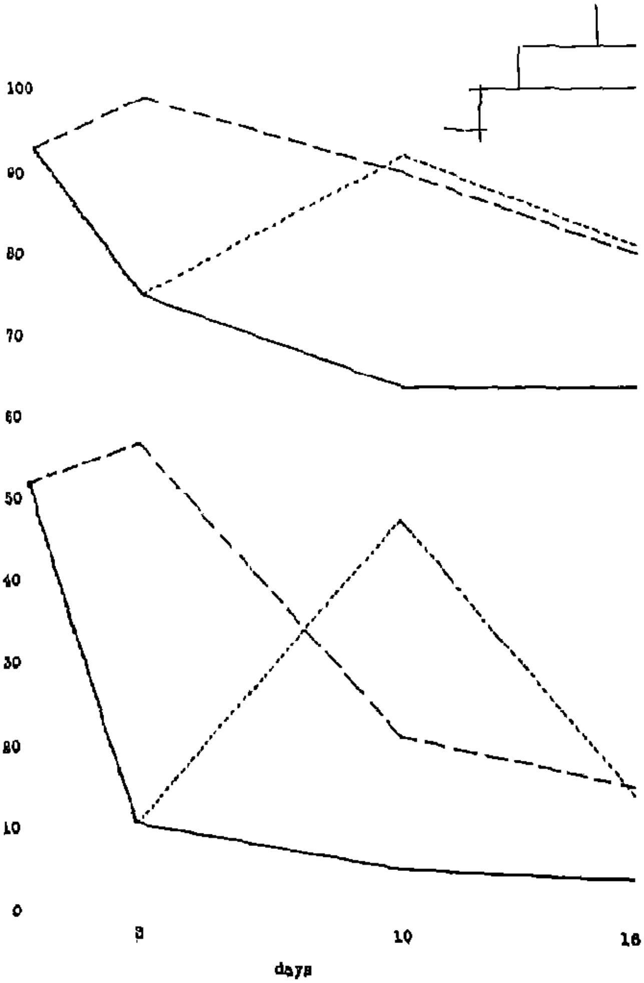
The same results are shown graphically in figure 2. The three squads are not graphed separately but only the average of all the squads—the data that appear in the bottom block of table 2. The three upper curves are for recognition and the three lower for recall. The abscissa represents time in days and the ordinate the per cent of retention. The heavy curve comprises the words that were presented only at the outset—the usual curve of forgetting. The dash line gives the retention for words with early follow-up. Of course its first ordinate must be the same as that for the ordinary curve of forgetting. After the follow-up, however, there is a rise in the curve. It subsequently drops but not as low as the original curve. The dotted line gives the retention for words with late follow-up. The first two ordinates must be the same as those of the ordinary curve of forgetting. After the follow-up there is an abrupt rise and a subsequent drop.

Inspection of the table and figure indicates clearly the advantage of the follow up. The curve rises after the follow-up and never drops as low as it does without follow-up. Moreover a superiority of the early follow-up to the late is manifest. The peak of the former curve is higher while there is little difference in the ordinates at the end.

Perhaps the best way to compare the two systems of follow-up as well as the ordinary curve is to total the four ordinates for each curve. This was done for each subject and then the results of all the subjects averaged. These averages in terms of per cent of possible maximum score appear in table 3.

Obviously the follow-up procedure adds distinctly to the

Figure 11



memory value. The significance of the differences from the standpoint of probable error is given in table 4. The difference between *E* and *I* for recall is 8.1 the probable error of difference while that between *L* and *I* is 6.1 P.E. The corresponding figures for recognition are 9.0 and 7.4. Moreover in recall all the total scores of the individual subjects are greater for *E* than for *I* and 98 per cent of the subjects have a higher total score in *L* than in *I*. With the recognition method the corresponding figures are 95 and 98 per cent.

TABLE 3

Total memory value (average of all subjects) in terms of per cent of maximum possible score

		RECALL	RECOGNITION
I	Initial only	19	74
E	Early follow up	37	92
L	Late follow up	31	86

TABLE 4

Differences divided by probable error and per cent of subjects yielding differences in the direction indicated

	E > *I*	*L* > *I*	*E* > *L*
D/P.E. recall	8.1	6.1	2.3
D/P.E. recognize	9.0	7.4	3.4
Per cent subjects recall	100	98	68 (15)
Per cent subjects recognize	95	98	88 (10)

The difference between *E* and *L* is not as marked as the above. In recall (table 3) the total memory value for *E* is 37 per cent as against 31 per cent for *L*, while with recognition the corresponding figures are 92 and 86 per cent. The former difference (table 4) is only 2.3 the probable error of the difference and the later 3.4 P.E. The distribution curves if plotted overlap considerably. However in recall 68 per cent of the subjects have a higher total for *E* than for *L* while 15 per cent (indicated in brackets in the table) have equal totals for *E* and *L*. This leaves only 17 per cent actually in the reverse direction from the general tendency. In recognition 88 per cent of the subjects have a higher total for *E* with 10 per cent equal. The preponderance of the evidence is thus in favor of the early follow-up.

SERIES III

Another series was performed similar to Series II except that the presentation was auditory rather than visual. The experimenter read the words aloud keeping time with a muffled metronome, set at one beat a second. The commodity was read on the first beat, the brand on the second followed by two beats pause. The next commodity was read on the fifth beat, etc. The same list was used as in Series II and the same mimeographed blanks. In the initial presentation the complete list was read twice. In the subsequent presentation of 30 words only one reading was given. A different group of subjects of course, participated—47 members of an elementary psychology class. Three squads were selected and the same schedule followed as in Series II. This group worked somewhat more rapidly than the others recording the results for recall and recognition in about 2 minutes and 1 minute respectively.

In scoring the results it was not possible to hold to absolute correctness of every letter in the recall tests because some of the fictitious names may have sounded differently from the way they were spelled. In scoring, a word was counted correct if the spelling indicated the correct sound.

The results are summarized in table 5 which is identical in form with table 2 and presented graphically in figure 3 which is identical in form with figure 2. The auditory data appears quite similar to the visual although the recognition curves are not as clear cut because, for some reason, the subjects scored higher in this series and in recognition came very near perfection.

The total memory value, i.e., the total of the ordinates of the curves is given in table 6 and the significance of the differ-

TABLE 5

Per cent of retention after different intervals with early and late follow up

SQUAD	RECALL					RECOGNITION				
	Presentation	0 days	3 days	10 days	17 days	0 days	3 days	10 days	17 days	
A	Initial	53	38	8	10	92	81	73	65	13
	Initial and 3 days		70	37	18		100	95	81	
	Initial and 10 days			55	27			98	95	
B	Initial	69	25	7	6	97	80	80	67	22
	Initial and 3 days		76	37	28		100	96	76	
	Initial and 10 days			59	27			97	75	
C	Initial	55	9	8	7	95	85	74	67	12
	Initial and 3 days		50	17	16		98	82	85	
	Initial and 10 days			50	21			96	79	
Average	Initial	59	24	8	8	95	82	76	66	
	Initial and 3 days		65	30	21		99	91	81	
	Initial and 10 days			55	25			97	83	

TABLE 6

Total memory value (average of all subjects) in terms of per cent of maximum possible score

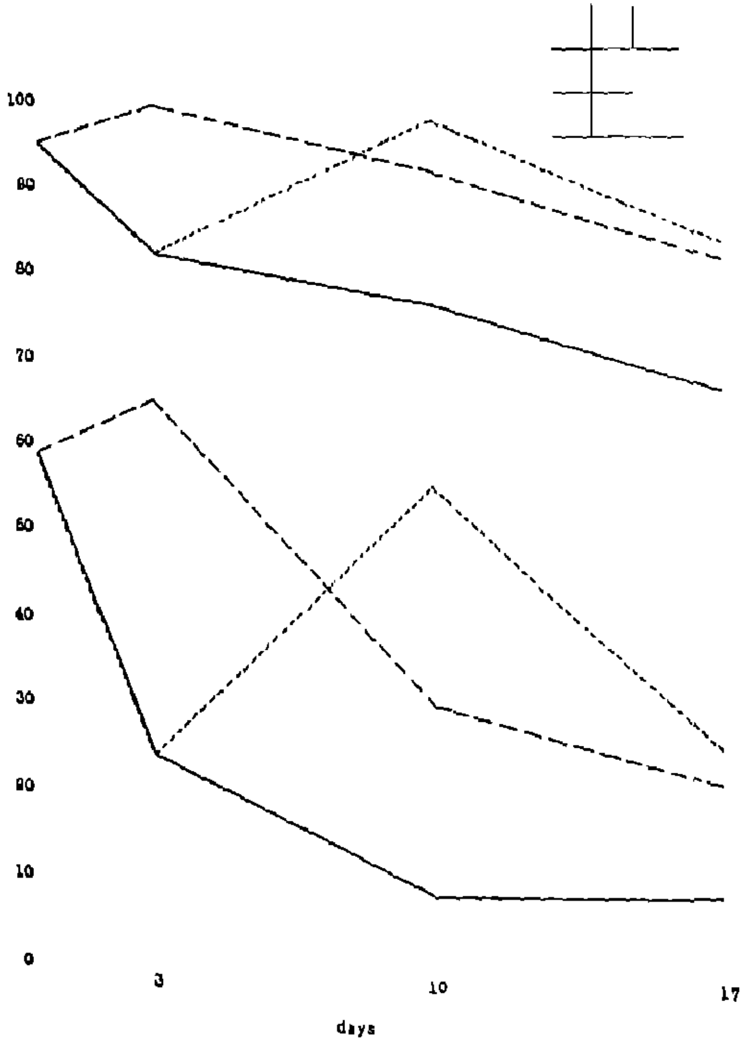
		RECALL	RECOGNITION
I	Initial only	25	81
E	Early follow up	44	93
L	Late follow up	41	90

TABLE 7

Differences divided by probable error and per cent of subjects yielding differences in the direction indicated

	$E > I$	$L > I$	$E > L$
D/P.E. recall	10.3	9.2	1.4
D/P.E. recognize	8.9	7.0	1.9
Per cent subjects recall	98 (2)	98 (2)	72 (4)
Per cent subjects recognize	98 (2)	91 (9)	66 (15)

Figure III



ences in table 7. These tables are identical in form with tables 3 and 4 respectively. The follow-up methods prove as superior to the other as they did in the preceding series. The advantage of the early over the late follow-up is likewise manifest although not quite as marked as in Series II. The differences are not as significant from the standpoint of probable error but the majority of the subjects have a higher score on *E* than on *L*. (The figures in brackets in table 7 indicate equal score.) The auditory series then confirms the visual.

SUMMARY

A study was made of the curve of forgetting with material designed to produce the kind of association frequently involved in actual advertising procedure. Pairs of words—a commodity and a fictitious trade name for it—were presented at constant rate. The subjects were tested for both recall and recognition of the trade names. Some of the words were tested immedately after the presentation and others after a lapse of time ranging from 3 days to 4 weeks. In some series a portion of the original list was presented a second time either 3 or 10 days after the initial presentation. It was thus possible to plot the ordinary curve of forgetting and that same curve as influenced by early or late follow up. Both visual and auditory presentation were employed.

The curve of forgetting with this sort of material shows the characteristic initial drop that has been found in other studies with other material. From the standpoint of total memory value the importance of the follow-up procedure is obvious. The early follow-up, under the conditions of the experiment, appears uniformly superior to the late follow-up. Conclusions drawn regarding advertising follow-up procedure on the basis of previous theoretical experiments would seem well founded.

SCHOOL ACHIEVEMENT AND SUCCESS IN LIFE

A. T. POFFENBERGER

Psychological Laboratory, Columbia University

Evidences of a positive relationship between success in school work and success in one's life work are interesting and important. A very close relationship, if such were found, between performance in college or technical school and in practical affairs would make one's school record a valuable diagnostic sign for vocational guidance work. It might become a valuable adjunct to tests or in some cases an adequate substitute for them. Other less tangible benefits from the discovery of such relationship might be found in a heightened respect for college work on the part of the student and his future employer.

We must depend upon the accumulation of studies of the degree of relationship in particular instances to provide the basis for evaluating school work as a vocational guidance device. One of the most recent and extensive investigations of this sort is that of Gambrill,[1] who compared college achievement in terms of scholarship and extra-curricular activities with income twelve and one-half years after graduation. Three of her conclusions will be cited here:

> 1. The relationship between college standing and income twelve and one-half years after graduation is too slight to warrant the use of marks as the chief basis for predicting the kind of success that the average employer has in mind when he consults the appointment committee in regard to students or graduates. Marks should, however, be considered as one factor in prognosticating vocational success, since the coefficients of correlation found were positive. The relative weight to be given scholarship in predicting later achievement in any given occupation can only be determined through further investigation.
>
> 2. The relation between extra-curricular activity and income is somewhat closer than that between scholarship and income. This emphasizes the necessity of faculty and administrative recognition of student

---

[1] B. L. Gambrill, College Achievement and Vocational Efficiency, Teachers College Publications, 1922.

> life as a selective agency of even greater importance than the curriculum so far as the qualities that make for future vocational success are concerned.
>
> 3. The results here reported prove nothing as to the relative contributions of college studies and of extra-curricular activities to the production of vocational success, nor as to the effects of study or neglect of study upon vocational success. Rather, they serve as measures of the two sets of activities, curricular and non-curricular, as selective agents in identifying the qualities which operate in later life to produce vocational success (page 78ff.).

Chapter 2 of this same study gives a good summary of the earlier work on the relation between school work and vocational success.

The most recent study of this type appeared in the *Military Engineer* for January-February, 1924, under the title, Relation of Class Record to Success. As this journal is not readily accessible to students of Applied Psychology, the author, Major Wm. E. R. Covell, Personnel Officer, Office Chief of Engineers, and Captain Lohr, Editor of the *Military Engineer* readily granted permission to reprint the article almost in its entirety. The principal changes are the elimination of two figures and the addition of a table of original data kindly furnished by Major Covell.

> Is high class standing in college an indication of future success? In the race of life, will the honor graduate maintain his relative standing ahead of the poorer student?
>
> In a course of psychology at the Massachusetts Institute of Technology, the text-books used deplored the lack of data to study the later results of class standing. This, together with an article entitled, "Are Honor Men the Best Men?" by Major J. G. Steese, Corps of Engineers, have induced an extended study of the records of the Military Academy at West Point.
>
> Here we have excellent materials for a laboratory test, for a record of the class standing at West Point of every graduate, and a history of his later life, may be found in the various editions of General Cullum's Register.
>
> We will extend this study to include the World War and will enlarge it by considering, in addition to purely military success, that in civil life. Also, in the hope that it may be interesting, there will be included the relation of class standing to athletics and to the dismissals from the army, the latter being the antithesis of success.

But before a start can be made in the study of class standing as a prediction of success, the latter term must be defined. What is meant by success? It would appear that some variation in its definition will not mar the value of any study made, provided the same definition remains fixed during the entire study. Because of this the definition used herein is arbitrary, but is applied to every graduate alike. We have taken as a test of military success the appointment of the graduate to the grade of Brigadier General or above, including the Brevet rank during the Civil War and the temporary rank during the World War. The rank of Brigadier General *in the Southern States during the Civil War has not been* included. But many graduates in this category are included in the civil list where the measure of success has been taken approximately as that required for inclusion in Who's Who, namely, bank president, member of Congress, mayor of city, editor, author, etc.

In the course of this study, the records of over four thousand graduates have been examined, from the class of 1818, when class standing was first introduced at West Point, to the class of 1905, which includes the last Brigadier General in point of graduation. With such a large number the law of averages is almost sure to work and the results obtained may be taken as a good indication of the general laws.

In the first place, it is desired to show the relation between class rank and success in after life. To accomplish this, the list of successful graduates has been arranged in the order of their relative class standing and this list divided into four numerically equal groups, as follows:

CLASSIFICATION	TOTAL NUMBER IN GROUPS	NUMBER SUCCESSFUL	PER CENT OF SUCCESSFUL
First quarter....................	1068	305	32.5
Second quarter.................	1068	251	26.7
Third quarter..................	1068	220	23.4
Fourth quarter.................	1068	164	17.4

But even more remarkable are the facts concerning the first, second, next to last and last man in each class, eighty-nine classes being available for study:

CLASSIFICATION	NUMBER SUCCESSFUL	NUMBER PER CENT	RATIO TO LAST MAN
First man......................	42	47.2	8.4
Second man....................	34	38.2	6.8
Next to last man..............	12	13.5	2.4
Last man.......................	5	5.6	1.0

The tables are self explanatory and show clearly and beyond argument that the higher a man stands in his class while at school or college the greater is his expectation of success in after life.

The following figures constitute the raw data from which the author constructs his briefer tables:

CLASS RANK	NUMBER OF MEN OF THIS RANK	TOTAL DISTINGUISHED	CLASS RANK	NUMBER OF MEN OF THIS RANK	TOTAL DISTINGUISHED
1	89	42	43	46	10
2	89	34	44	44	7
3	89	31	45	41	6
4	89	34	46	39	4
5	89	24	47	39	6
6	89	25	48	38	7
7	89	20	49	35	5
8	89	26	50	34	4
9	89	21	51	33	5
10	89	21	52	28	5
11	89	26	53	27	6
12	89	20	54	22	5
13	89	19	55	22	3
14	89	24	56	20	4
15	89	16	57	19	1
16	89	16	58	18	3
17	89	28	59	16	2
18	89	26	60	16	0
19	89	21	61	16	2
20	89	20	62	15	3
21	89	18	63	14	2
22	89	23	64	13	1
23	88	21	65	12	1
24	87	18	66	12	0
25	85	18	67	10	2
26	85	12	68	9	0
27	83	14	69	9	1
28	81	17	70	9	1
29	81	17	71	9	0
30	80	22	72	8	0
31	78	20	73	7	0
32	78	24	74	6	0
33	76	18	75	6	1
34	74	13	76	5	0
35	73	8	77	4	0
36	72	12	92	4	0
37	68	16	93	3	0
38	65	9	113	3	0
39	62	9	114	2	
40	61	10	123	2	
41	54	10	124	1	
42	51	8			

Ignominious dismissal from a man's chosen profession is surely the antithesis of success. What then, does the relation between class rank and dismissals from the Army show? This is a little more difficult question as it depends more on a man's morality than on his ability. Also, it is necessary to consider this question from two angles, according, perhaps, to the point of view of the reader. Because at the time of the Civil War, while practically all of the southern officers resigned, fourteen were dismissed from the Army of the United States for "treasonable utterances." These have not been included in the table:

CLASSIFICATION	TOTAL IN GROUP	NUMBER DISMISSED	PER CENT DISMISSED
First quarter	1,068	11	18
Second quarter	1,068	10	16
Third quarter	1,068	17	27
Fourth quarter	1,068	24	39
Total	4,272	62	100

The numbers considered here are too small to establish any general law, but it is interesting to note that the tendency seems to be to present a table of percentages just the reverse of the table shown for successful graduates. The latter gives for the upper and lower halves of the class:

Successful men	per cent
Upper half of class	59
Lower half of class	41

while the former gives:

Dismissed	per cent
Upper half	34
Lower half	66

The question of the first man and last man in this case is debatable because of the small numbers, but the facts may be mentioned:

Dismissed

	per cent
First man in class	2
Last man in class	3

The perpetual argument as to the effect of athletics on class standing has waged for years but, for the most part, with few available facts. For this reason, it may be of interest to cite the results of this study and to comment thereon.

In this portion of the study the test of athletic prowess has been taken as the winning of the coveted "A" while at the Military Academy. The classes from 1891, when these marks of athletic distinction were first awarded, to include the class of 1920 have been taken. The results are as follows:

CLASSIFICATION	AWARDED "A"	PER CENT
First quarter	83	20.9
Second quarter	107	26.0
Third quarter	105	20.3
Fourth quarter	102	25.9
Total	307	100.0

The quarters are very well grouped except the first and this perhaps should be discussed. One explanation may be that participation in athletics tends to lower marks received so that, though the athletes may have been evenly divided at first, their athletic training caused a few to drop from the first quarter to the second, from the 2nd to the 3d, from the 3d to the 4th, and from the 4th quarter to a non-graduated status. The 102 of the 4th quarter includes 13 men who won their "A's" but never graduated or, it may be that high standing in class tends to lead a man into other branches of academy activity. But, this question can hardly be considered as settled, and offers an opening for future study.

A more interesting question is "what effect does the winning of the college letter have on success in after life?" In this case, too, the numbers considered can hardly be said to give conclusive results, but they are most remarkable and interesting.

The classes from 1891 to 1905, inclusive, have been taken for study, the former being the first class in which "A's" were awarded and the latter the last class to produce a man of distinction.

	NUMBER	DISTINGUISHED	PER CENT
Total in classes..................	1,068	125	12
Total "A" men.................	174	21	12

In other words, the athlete had exactly the same chance to become distinguished as the other men of his class and no more.

Our conclusions will be brief and simple and, except perhaps, for the first, open to the objection that they have not been completely proved. However, the data in this case certainly tend to show:

1. That the higher a man stands in college, the greater is his expectancy of success in after life.

2. That a man in the lower part of the class is more apt to become a moral failure than a man in the upper part.

3. That class standing is no criterion of athletic ability.

4. That an athlete has the same expectancy of success and no more than any other man of his class.

THE PRESENT STATUS OF THE WILL-TEMPERAMENT TESTS[1]

MARK A. MAY

Teachers College, Columbia University

The WT[2] tests are one of the best instances we have of an attempt to measure temperament objectively. The underlying conception is that fundamental behavior patterns are revealed in certain skilled motor performances such as speech and handwriting. Downey (9, p. 49) asserts that these tests are the results of several years of inquiry into the nature of muscle-reading and handwriting. Her experimental studies of muscle-reading revealed the fact that certain forms of mental content are expressed in subtle muscle twitches, and also in delicate expansions and contractions. There are great individual differences in this phenomenon. Some persons' muscles are easier to read than others because (she thinks) they are more explosive and have fewer inhibitions. In some persons there seems to be an easy and forceful flow of nerve currents into the muscles; in others, the flow seems obstructed. This immediately suggests temperament types. Thus she says (9, p. 54): "Temperament appears to be very determined by the readiness with which the motor discharge, which initiates movement, occurs in the nervous system and the degree to which it stimulates consciousness."

Downey found that good subjects for muscle-reading demonstrations could be selected by a simple handwriting test. Persons who write more rapidly under distractions are usually good subjects. This very fact together with the ancient belief that character is revealed in handwriting led her to investigate graphology scientifically. These studies resulted in a denial of most of the claims of the graphologists but led her straight to the WT tests. The central idea of the tests is that individuals vary in the amounts of nervous energy[3] they possess and also in the ease or difficulty with which it is discharged into the motor pathways. Temperaments can be thought of as arranged on a linear scale. At one end is the individual whose activity is free, forceful and explosive; at the other end is the person whose energy is limited and inhibited. In the middle are those with varying amounts of energy and varying degrees of inhibition. Temperament is then a function of two things, *amount* of energy and its *inhibition* or *facilitation*.

---

[1] This is one of a series of papers on A Survey and Evaluation of Existing Character Tests being prepared in connection with an inquiry in character education made possible by a grant to Teachers College from the Institute of Social and Religious Research. MS. submitted October, 1924.

[2] WT will be used throughout this paper as an abbreviation for Will-Temperament.

Richardson (17), an English school master, advanced practically the same theory of character in 1915. His notion is that character is based fundamentally on an inherited "responsiveness" to stimuli and that individuals vary in the nature, magnitude, and speed of their responses. Character varies with conditions that cause variations in the nerve impulses. These variables are nutrition, temperature, fatigue, concentration of solution, pressure and other physical and chemical agents. The two fundamental measures of character are *responsiveness* and *inhibitions*.

Downey attempts to measure temperament by measuring (1) the speed of reactions, (2) the forcefulness and decisiveness of reactions and (3) the persistence or perseveration and care with which the reactions are made. These three variables seem to her to reveal the amount of available energy. For each of these she has devised four tests.

1. "Speed and fluidity of reaction" is measured by the following tests:

A. Speed of movement. Measured by the average time taken in two trials to write the phrase "United States of

---

[3] Unfortunately she does not define "nervous energy."

America." The subject is instructed to write at his *usual* speed and in his *usual* style.

B. Freedom from load. Measured by the ratio of natural speed of handwriting to fastest speed.

C. Flexibility. Measured in two ways: First, by the degree of success in disguising his handwriting; second, by speed and success in copying a model. The phrase written is the same as above.

D. Speed of decision. Measured by the time taken to check certain character traits which the subject thinks he possesses.

2. Forcefulness and decisiveness of reaction, "the aggressive traits," measured by the following:

E. Motor impulsion. Measured by writing signature in four ways. (*a*) Usual style and speed with eyes closed. (*b*) Eyes open but counting by 3's aloud. (*c*) Eyes closed but counting by 3's aloud. (*d*) Eyes closed, counting by 2's, examiner taps with pencil. Score: Changes in size and speed of writing under distraction.

F. Reaction to contradiction: Suggestibility. Early in the examination S is asked to choose between two envelopes marked M and N. He is told that one contains an easy test and the other a hard one. This test consists in trying later to convince him that he chose the one he did not. The score is the degree to which he sticks to his original decision.

G. Resistance to opposition. Tested by blocking the movement of the pencil when S is writing blindfolded. Score is the exertion he exhibits in overcoming the obstacle.

H. Finality of judgment. Measured by returning to the Speed of Decision test and asking him to revise his judgments of the traits checked. Score is the time spent in revision and the number of changes made.

3. "Carefulness and persistence of reaction" is tested as follows:

I. Motor inhibition. Measured by the time of retarded handwriting. The subject is asked to write the phrase "United States of America" as slowly as he can. Score is the average of three trials.

J. Interest in detail. Measured by the exactness with which a model is copied, first when he is instructed to be as exact as possible and second when he is told to do as well as possible and choose his own speed. Score is the relative time taken and the exactness achieved.

K. Coördination of impulses. Measured by the degree of success achieved and the time taken in writing the phrase in a very narrow space.

L. Volitional perseveration. Measured by the time taken in working out a disguised hand. The subject is told to write the stock phrase in such a disguised way that even an expert could not detect it.

We now have before us the WT tests in their historical setting. There are several questions to be asked about them.

1. Are these handwriting exercises adequate measures of such things as speed of reaction, forcefulness of reaction, etc.? Of the 12 WT tests 9 are handwriting tests. Downey faces this problem and in her book (9) takes each test separately and attempts to justify its existence by showing that the handwriting exercise involved is a true measure of the trait in question. For each test she presents practically the same arguments, which are in the main as follows:

a. She appeals to her previous experimental work on graphology and muscle-reading. Of the "Speed of Movement" she writes (9, p. 87): "The most serious question that can be raised concerning the use of graphic speed in the test is that which asks whether rapidity of handwriting is, in any way, an index of general speed of movement. In another connection the author has shown that graphic speed is, in fact, symptomatic of general bodily speed of movement." Here she refers to her book on "Graphology and the Psychology of Handwriting" Chapter VII.

Turning to this reference we find reported an experiment in which six student judges and five faculty judges rated 12 students on a list of movement traits, one of which was "rapid or slow." She admits that the judges "reported great difficulty in passing these judgments." These ratings were correlated with her handwriting records and the correlation between ratings and normal writing was 0.55 and with speeded writing 0.61. Assuming that these judgments are entirely reliable, one could scarcely conclude that "graphic speed is, in fact, symptomatic of general bodily speed" in the face of a correlation of 0.61. Before such a conclusion can be drawn speed of handwriting should be shown to correlate highly (0.80 or more) with such bodily functions as tapping, reaction-time and other measures of bodily movement.

b. As a second criterion she uses inter-correlation amongst the tests of a group. That is, if the first four tests all measure some form of speed of reaction, there should be a high inter-correlation between them. She says that Speed of Movement, for example, correlates positively "with all other items in the speed group, which can only be interpreted as meaning that it has characterial significance." In the first place, Ruch and Del Manzo (19) have shown that these inter-correlations amongst the tests of a group are not always even positive, and even if they were all very high, they would not prove her point. Here are four tests all purporting to measure some aspect of a general factor called "Speed and Fluidity of Reaction." If there is a high inter-correlation among these four traits it means only that they measure the same thing, but it does not tell what they measure.

c. Another criterion is a set of correlations between the Downey individual form and the Carnegie Group Form reported by Ream (16). These correlations will be examined later. The point here is that a high correlation between the group form of a test and the individual form does not prove or even hint that the test measures the trait suggested by the label. However, in fairness it should be said that she simply reports these inter-correlations without urging that they are significant.

d. Then she makes a strong appeal to personal observation and to knowledge of personal acquaintances. Of Freedom from Load she writes: "That the ratio of speeded to normal writing-times is actually an index of amount of general inertia or 'load' appears highly plausible from a study of the records of individuals with whom I am intimately acquainted." Of course, there can be no objection to any investigator making use of all possible sources of information, but the appeal to personal experiences needs to be checked by laboratory experiments.

Downey (9, p. 173) admits the general criticism that these handwriting exercises are probably not true measures of the traits suggested in their labels. She recognizes the fact that the more scientific procedure would be to label them "Disguised Handwriting Test," or "Speed of Handwriting Test," etc., instead of the characterial labels they now bear. But her excuse is that the present procedure has aroused a degree of interest among psychologists that the more scientific procedure would not have aroused. This is undoubtedly true.

2. What is the reliability of these tests? The usual method of determining reliability is by finding the self-correlation of tests. The data are obtained by giving the same test or duplicate forms twice or more. We have a few instances of this with the WT tests. Ream (16) gave the Downey individual form and the Carnegie Modified Group Form to 23 subjects. Meier (15) gave the Downey individual form and the Downey group form to 64 subjects. Their correlations are:

	TRAITS[*]					
	A	B	C	D	E	F
Ream	0.72	0.05	0.12	0.42	0.50	0.42
Meier	0.57	0.06	−0.01	0.18	−0.23	−0.22
	G	H	I	J	K	L
Ream	[†]	[†]	0.55[‡]	0.72[‡]	0.10	0.90[‡]
Meier	−0.05	0.65	0.55	0.48	0.04	0.20

Meier found the correlation between total scores to be 0.60.

[*] The letters refer to the traits as lettered earlier in this paper.

[†] No equivalent.

[‡] Correlation ratio, but he does not say which one.

The above figures are not quite fair since neither the Carnegie group test nor the Downey group form are the exact equivalent of the Downey individual form. Hence, these low self-correlations mean only that, on the whole, the individual Downey tests and their supposed group equivalents do not measure the same thing.

Ruch and Del Manzo (19) attempted to measure the reliability in a different manner. According to Downey, tests A–D, that is the tests of the speed group, reveal the "hair-trigger" type of personality; tests E–H reveal the "wilful aggressive" type; and tests J–L show a "slow accurate, tenacious," pattern. If this grouping of tests is correct one would expect the following things to be true:

a. That the *inter-correlations among* the tests of a group would be positive and high.

b. That any one test of a group would correlate positive and high with the total of the remaining three tests in that group.

c. That the correlations between the tests of one pattern and those of another pattern would be low.

We have two tests of the first proposition. Downey (9) plotted scatter-diagrams for each combination of the tests. She did not compute the coefficients of correlation but calculated the per cent of agreement within three places and compared this with chance agreement which she says is 58 per cent. She interprets an agreement of 75 per cent within three places as a high correlation. Unfortunately she does not report all of these. I have collected by patterns or groups those reported.

PATTERN I—TRAITS A–D		PATTERN II—TRAITS E–H		PATTERN III—TRAITS J–L	
	per cent		*per cent*		*per cent*
AB	75	EF	65	IJ	77
AC	67	EG	66	IK	65
AD	69	EH	73	JI	74
BC	+	FE	66	JK	+
BA	76	FG	69	KJ	69
CD	+	FH	65	LJ	72
CA	+	GF	+		
CB	+	HE	79		
DB	64	HF	65		

These are of the nature of correlation ratios. Traits AB, for example exhibit agreement of 75 per cent within three places when B is compared with A, and 76 per cent when A is compared with B. In the above scheme the first letter stands for the independent trait. The (+) plus sign means that she simply reported a positive relationship without stating the amount.

Ruch and Del Manzo (19) report inter-correlations on 146 high school students. Their results by patterns are:

PATTERN I		PATTERN II		PATTERN III	
	r		*r*		*r*
AB	0.36	EF	−0.05	IJ	0.08
AC	0.28	EG	0.00	IK	0.12
AD	0.37	EH	−0.02	IL	0.00
BC	−0.05	FG	−0.16	JK	0.24
BD	0.33	FH	−0.15	JL	0.18
CD	0.05	GH	0.01	KL	−0.18
Average	0.22	Average	−0.06	Average	0.07

They also obtained the correlation between each test in a pattern and the sum of the three remaining tests.

(A	B	C	D)
0.54	0.28	0.10	0.30

(E	F	G	H)
− 0.08	− 0.23	− 0.13	− 0.09

(I	J	K	L)
0.07	0.33	0.07	0.05

This means that trait A correlated 0.54 with B + C + D, that B correlated 0.28 with A + C + D, etc.

Ruch and Del Manzo attempted to find the coefficient of reliability of the entire test by using the Brown formula for correlation of alternate tests. Since there are three main patterns each tested in four ways, two half-scales can be made in three different ways.

1. The correlation of odd numbered with even numbered tests: $r = 0.17$, $B = 0.29$.

2. The first two tests in each pattern with the last two: $r = 0.31$, $B = 0.47$.

3. The 1st and last test in each pattern with the middle two: $r = 0.24$, $B = 0.38$.

B is the correlation, prophesied by the Brown formula, that *would result if two similar whole forms were correlated.*

These low reliability figures may mean either that the WT tests are rather unreliable or else the method of determining their reliability is wrong.

3. What is the validity of the W T tests? Against what criteria have they been checked?

a. Validation by checking against ratings. Ruch (18) secured independent ratings in the 12 Downey traits from two groups of judges (a faculty group and a student group) on 20 advanced students. Scores in the twelve traits were correlated with the ratings for each subject. This yielded twenty correlations for the pooled estimates of student judges, and twenty for the faculty judges. The student estimates and scores correlated all the way from −0.64 for one subject to +0.54 for another, with an average of 0.08; test scores and faculty judgments correlated from −0.18 to +0.29, average 0.01; pooled student judgments with pooled faculty judgments correlated from −0.65 to +0.88 with average at +0.40. Another set of correlations was made by traits, the subjects being the variables, and this yielded twelve coefficients, one for each trait. They are:

Test with faculty estimates run from −0.33 to +0.51, average +0.07.
Test with student estimates run from −0.23 to +0.53, average +0.20.
Faculty estimates with student estimates +0.22 to +0.86, average +0.62.

Ruch and Del Manzo (19) continued this study using as subjects 146 high school students. In order to avoid some of the errors involved in rating schemes they asked each teacher to pick out of each class ten students strongest in "will-power" and ten who are weakest in this respect. The teachers agreed almost perfectly on 28 students out of the 140; they were about evenly divided between the sexes and between the "strongs" and the "weaks." This group of 28 were given the Downey tests and were rated on the Downey traits according to the descriptions of the traits taken verbatim from her manual. The correlations between test score and the pooled ratings are:

Trait	A	B	C	D	E	F
r	0.37	0.02	0.30	− 0.29	0.10	− 0.03
Trait	G	H	I	J	K	L
r	− 0.04	0.17	− 0.09	0.51	0.53	0.17

Meier (15) gave the Downey individual tests to 106 high school students, sexes evenly divided, ages running from 13 to 20. Three judges rated each pupil for the twelve Downey traits on a "man-to-man" type of rating scale. The judges in each case were (*a*) the teacher who knew him best, (*b*) a parent, and (*c*) a friend over eighteen who had known him for at least two years. Meier took extreme pains to get the descriptions of the traits clear. He even sent them to Downey who approved them with slight modifications. His correlations are:

Trait	A	B	C	D	E	F
r	0.21	0.07	0.11	0.19	0.10	0.05
Trait	G	H	I	J	K	L
r	0.14	0.14	0.24	0.21	0.07	0.03

Herskowitz (13) selected three groups of persons who were well acquainted with each other within the group. They rated each other on each of the Downey traits using a scale of 0 to 10. There were 7, 8 and 9 persons respectively in the three groups. He describes the personnel of the three groups giving the age, sex, schooling, occupation of each individual and also states the relationships existing among the members of each group. First he pooled all estimates and all scores and correlated 2152 independent judgments with that many scores and obtained a correlation of 0.06 P.E. 0.012. Believing that this low correlation was partly due to heterogeneity of material he next correlated the average of each individual with score in traits which yielded a correlation of 0.13, P.E. 0.038. Next he correlated the scores in the test with average judgments for each test taken separately. The results are:

Trait	A	B	C	D	E	F
r	– 0.03	0.18	0.12	– 0.05	0.21	0.10
Trait	G	H	I	J	K	L
r	0.18	0.35	0.29	– 0.001	0.28	– 0.14

He undertook to correct these correlations for attenuation by a method which need not be discussed here but which is statistically sound. He finds that under the most favorable conditions one could not expect a correlation between Downey test scores and ratings to be higher than 0.40. He concludes that the Downey tests in their present form could not be used profitably in his studies of racial temperaments.

Miss Bryant (1) gave the Downey tests to one hundred delinquent boys at the Whittier State School. She secured estimates of each boy from three competent judges. They judged only the last group (i.e., the volitional traits) or rather traits corresponding to the last group. The traits judged were self-confidence, self-control, power of concentration, persistence, initiative in every day affairs. The correlation between test score and ratings for these traits is 0.29.

The results of all these attempts at validation by the rating method are uniformly ambiguous. Nearly all the correlations are low. Does this mean that the tests are not valid, or the ratings unreliable, or that this is not the proper method of testing the tests? An examination of the data reveals the fact that these ratings are about as reliable as ratings usually run. Other kinds of tests have been validated by this method. We venture the assertion that after ample allowance has been made for errors in ratings the correlations will still be low.

One difficulty seems to be that judges simply cannot identify the psychological traits supposedly measured by the Downey tests. At least, they cannot identify them from Downey's descriptions of them. For example, take the first trait, Speed of Movement. The rater is asked to judge an individual on his general speed of movement. The judge naturally thinks about how quickly he moves, how fast he walks, how fast he talks, etc. But the test measures how fast he writes. If the judges were asked for an opinion of the subject's speed of handwriting, the correlations might be higher. If this is the case, then these results show that the Downey tests do not measure the traits named in the labels.

Another possibility is that since the Downey tests attempt to measure a very subtle relation between the amount and availability of energy and how it is expressed, the average judge cannot evaluate this for any individual. This would mean that the Downey traits are simply too subtle to permit of judgments and hence the test cannot be validated by the rating method. This seems to be Downey's opinion. But on the other hand, she does not discard ratings entirely. In discussing the appropriateness of a certain handwriting exercise as a measure of a trait she continually appeals to her judgment of "personal acquaintances."

In her book (9, p. 217) she uses a self-rating scheme as a means of validation. She asked several "highly educated individuals" to rate themselves on a scale of ten points for each of the tested traits. In each case a comparison was made between the self-rated "profile" and the test-score "profile." She presents three of these double profiles all of which show a very close correspondence. Next she secured similar self-ratings from 21 judges "all of superior education and intelligence" and calculated the average amount of displacement of ratings from test score for each trait. Her results are

TRAIT	AVERAGE DISPLACEMENT	TRAIT	AVERAGE DISPLACEMENT
A	1.45	G	1.31
B	2.02	H	2.71
C	1.07	I	1.52
D	1.65	J	2.23
E	1.17	K	2.85
F	1.53	L	2.21

I take it that this table means that the average difference between test-score and self-rating on trait A is 1.45, etc. This statistical procedure is ambiguous. It would have helped in the interpretation if she had published some measure of the variability of these differences. Their meaning is illuminated somewhat when we remember that each scale contains ten points. Hence the above with the decimal moved *one place to the right will tell the percentage of the scale.* For instance, the average deviation in trait K is 28.5 per cent of the total scale. In this particular trait I would judge that the correlation is zero or negative. So even with "highly educated and intelligent judges" each judging their own traits, the rating method of validation is not wholly satisfactory.

One wonders what would happen if a group of eminent psychologists, skilled in the art of introspection and possessed with keen judgment and insight into the subtle workings of the minds of others, should make a careful study of the fundamental ideas back of the WT tests so as to become thoroughly saturated with them, and then try their hands at rating themselves and their intimate acquaintances.

b. Validation by correlation with other objective tests. We have two instances of other tests devised to measure the WT traits. Downey and Wagoner (22) worked out a will-temperament speech test. The idea was to parallel the handwriting tests with similar speech tests. For instance, the speed of movement is measured by the speed of reciting the alphabet three times without stopping. The correlations between the handwriting tests and their speech equivalents are:

Speed of movement	0.05
Freedom from load	0.20
Flexibility	0.39
Motor Impulsion	0.30
Reaction to contradiction	0.60
Motor Inhibition	0.63
Interest in detail	0.38
Coördination of impulses	0.11
Volitional perseveration	0.46

Miss Downey attributes some of the low correlations to attenuating conditions which she enumerates. She does not offer any corrections for attenuation.

Filter (11) studied two of the Downey traits rather extensively by checking them against ratings and also against other objective tests designed to measure the same traits. Unfortunately he selected two traits which are not based on handwriting. They were "reaction to contradiction" which he classifies as "self-assurance" and "speed of decision." He checked the "reaction to contradiction test" (the envelope test) against two other objective tests. (1) A String Figures test which calls for an estimate of one's capacity to reproduce on his fingers with a string a printed figure or diagram. (2) The Découpage test (about the same as the Binet paper folding test) which calls for an estimate of one's ability to reproduce a visualized diagram from a folded and cut paper. The correlation between the Downey test and each of these two tests is −0.10. He validated by ratings all three of the tests (the Downey and his two) and found the Downey test much lower than his own.

He selected and devised six speed of decision tests including the Downey character trait checking test and five others. He gave these tests to 142 subjects. The correlations between the Downey test and the other five are 0.43, 0.47, 0.56, 0.31, 0.51, respectively. Further analysis of his results revealed the fact that only about 10 per cent of his subjects were uniformly high on all tests; 8 per cent constantly mediocre, and 13 per cent constantly low. The remaining 69 per cent are neither beast, bird nor fish in speed of decision.

c. Validation by identification of profiles. Downey (5) instructed a group of judges in the meaning of the twelve traits. Each judge was presented with twelve profiles and a list of the names of the persons to whom they belonged. The task was to match the name with the profile.

Correct identification of profiles ran from 0 to 5 out of 12, or from total failure to identify any profile (1 judge) to 41 per cent of successful identifications (2 judges). The percentage of successes for the total of 144 judgments (12 judgments by each of 12 judges) was 22, where chance success would be less than 1 per cent.

As these results were somewhat unsatisfactory she tried it again, this time submitting profiles in groups of three, asking each judge to chose from the three one profile that best fitted a given person. In series A the three profiles were similar; in series B they represented contrasting types. There were in series A 69 faculty judges, 54 men student judges and 37 women student judges. Their average successes were 33.3, 44.4 and 51.3 per cent correct. (Chance success, is, of course, 33.3 per cent.) In series B, the average successes of the three groups of judges were 78.7, 71.9 and 58.6 per cent, respectively. Again she does not publish a measure of variability but gives the range as from 0 to 100 per cent correct. She examined in detail certain of the cases of mistaken identity and found in most cases the reason for failure. She concludes that in these cases the judge was judging some superficial appearance rather than the underlying temperament traits.

d. Validation by differentiation. Do the tests have differential value? Will they distinguish between the delinquent and the non-delinquent, between the negroes and whites, between the successful and the unsuccessful in salesmanship or other affairs?

In 1918 Downey studied insane and psychopathic individuals with the WT tests. She found what she thinks is a "characteristic profile" for dementia praecox, manic, depressed, and psychopathic patients. But these groups studied were all very small (four or five each) with considerable variability within each group.

Bryant (2) compared a group of 100 delinquent boys at the Whittier State School with a group of 100 non-delinquent boys selected in such a way as to be comparable with the delinquents in age, intelligence, etc. She compared the total scores of these two groups on the WT tests.

	RANGE	MEDIAN
Delinquents	22–75	45.5
Non-delinquents	40-81	61.0

The differences between the medians is 15.5 points which seems reliable, although she does not report a measure of its unreliability. Assuming that the S.D. of the total is 20 points (one-sixth of the range) then the P.E. of the difference is about 2.5. This is verified somewhat by other comparisons. From her graphs I would judge that not more than 15 per cent of the delinquents reached the median of the non-delinquents. She states that in the delinquent group 68 per cent of the cases are within the 40 to 50 quintile; while in the control group 67.5 per cent of the cases are in the 50 to 69 quintile.

She next compared the delinquents with the control group trait by trait. The traits in which they differed most are speed of decision, motor inhibition, coördination of impulses. In the coördination of impulses the average delinquent score was 1 point, the average control score was 7 points out of 10. Again she does not report the unreliabilities of these differences. She did, however, compute the correlation between each test and total score for each group, and compared the groups in this way. But such differences in correlations are hard to interpret.

Freyd (12) gave the Carnegie group form of the WT tests to groups of socially inclined and mechanically inclined individuals. He had two mechanical groups, mostly engineering students, and three social groups, all salesmen or students of salesmanship. Here is his summary table:

TRAIT	SALES GROUP AVERAGE	MECHANICAL GROUP AVERAGE	DIFFERENCE	RELIABILITY
Speed of decision (1)	24.34	28.44	4.10	∞
Speed of decision (2)	21.63	25.66	4.03	555,556
Flexibility	3.00	2.60	0.40	2,301
Motor Impulsion	197.98	201.64	3.66	8
Reaction to Contradiction	17.10	25.59	8.49	∞
Ratio of 1 and 2 above	122.00	125.76	3.76	5.3
Motor inhibition	14.18	9.00	5.18	31,546
Coordination of impulses	73.62	94.73	21.11	∞
Interest in detail	9.68	6.75	2.93	∞
Volitional perseveration	103.58	101.76	1.82	3.2

(1) Speed of decision in checking "better traits."

(2) Speed in checking "personal traits."

The last column states the chances that a real difference does not exist. For instance the chances are 1 to 8 that there is no real difference between the groups in "motor impulsion."

McFadden and Dashiell (14) sought differences between whites and negroes with the WT tests. They compared a group of white and negro high school and college students. The two groups were not exactly comparable because the negroes were older for their educational status than the whites. There were 78 in each group. Their main results are:

TRAIT	WHITE AVERAGE	NEGRO AVERAGE	DIFFERENCE
A	5.40	3.60	1.80
B	6.06	5.76	0.30
C	4.65	4.11	0.54
D	3.66	4.28	0.62
E	4.14	5.31	1.17
F	5.65	6.45	0.80
G	5.63	5.19	0.44
H	5.23	5.52	0.27
I	5.01	5.28	0.27
J	6.80	0.67	0.22
K	4.82	3.01	1.81
L	6.23	5.74	0.49

The authors do not report the unreliability of these differences but they do give figures from which they may be computed for the college group. I find that these P.E.'s of differences run from about 0.30 of a point to 0.50 of a point. In the above table there are not over three reliable differences. The traits in which these differences exist are A, E and K which are speed of movement, motor impulsion, and coördination of impulses, the whites excelling in the first and last.

The total scores for the two college groups were compared. They are: The mean total score of the white college group is 68.08, of the negro college group 59.05; the medians are 70 and 57. There is a difference of about 9 points between the means and 13 points between the medians. The P.E. of these differences is not given, but I judge it to be about 1.5 for the mean and about 2.0 for the median. Only 15 per cent of the college negroes passed the median of the college whites.

In these studies two results seem to stand out clearly. First, there is no characteristic or easily identifiable profile for any one group. No one could look at a group profile and say whether it is that of a group of whites, negroes, delinquents, normals, insane or anything else. Second, total scores apparently have some differential value.

e. Validation by testing the prognostic value of these tests. The question is can these tests be used for predicting any kind of human behavior?

Clark (3) scored the conduct-record of delinquent boys at the Whittier State School on a scale of 1, 2, 3, 4, 5 according as it was reported bad, poor, fair, good, excellent. The correlation between the average response rating and the total score on the WT tests is -0.22. But the correlation between WT total score and *change* in conduct-response is $+0.26$. Improvement in response may be predicted to a slight degree from the WT score.

Poffenberger and Carpenter (20) studied the school success of 97 children in grades 6, 7, and 8, ages 10:11 to 15:7, I.Q.'s, 82 to 160. Each pupil was rated for "school success" by three teachers and the ratings were averaged. The average ratings were correlated with intelligence test scores, 0.70, 0.74, 0.69, for grades 6, 7 and 8 respectively; and they correlated with I.Q. 0.70, 0.77, 0.77 for the three grades.

From the scatter-diagrams two groups were selected. The "success" group with a higher rating for school success than I.Q. would indicate, the average I.Q. being 106.6. The "Failure" group with lower success rating than I.Q. would indicate. The average I.Q. of this group was 116. Each group was tested with the WT tests. The question is, are there any temperament traits or groups of traits that will differentially characterize the success group and the failure group. There were certain traits in which the success group always stood high and the failure group low. These were interest in detail, assurance, volitional perseveration. The success group was low in freedom from load and motor inhibition. By pairing each of the twelve groups in combinations and studying the scatter-diagrams, it was found that the success group possessed certain combinations of traits. Thus, there are two factors really, one is the trait and the other is its relation with other traits. For instance, speed of decision when coupled with coördination of impulses makes for success; but when coupled with low coördination or lack of coördination makes for failure. By some such scheme as this, it is possible to predict school success or failure by a combination of intelligence and temperament tests.

Stone (21) used a modified form of the WT test in studying *the disparity between the intelligence and scholarship records* of Dartmouth students. He used the ratio between normal and speeded handwriting as a measure of *tension* and the ratio of normal to retarded handwriting as a measure of perseveration. He believes that the discrepancies between intelligence and scholarship may be accounted for by some such tests as these.

Downey (8) has made two or three similar studies the results of which are summarized in her book (9, ch. 17). This summary consists mostly of composite profiles and case studies which I cannot reproduce here. She summarizes her results:

> *Students with high intelligence rating but a low WT score* usually succeed academically, although not to the degree warranted by their intelligence, and, with a certain combination of traits may fail. Students with a low will-score who are rated in either group III or IV[4] for intelligence, fail to make good as students. Students of inferior intelligence but strong temperament qualities may succeed in maintaining a passing grade.

She found a correlation of 0.40 between intelligence scores on the Thorndike Examination and grades; and a correlation of 0.32 between WT scores and grade; and a correlation of 0.554 between a combination of the intelligence and the WT scores and grades. As supplements to intelligence tests the WT tests have great promise. But what behavior may be predicted by the WT tests alone has not been determined.

---

[4] Refers to Thorndike's grouping and means average or low intelligence.

4. What is the relation of the WT tests to intelligence? The following correlations have been reported:

INVESTIGATOR	TEST	SUBJECTS	r
Downey	Binet (Terman)	20 high school students	0.77
Downey	Thorndike	34 high school students	0.47
Downey	Alpha	"A more uniform group"	0.20
Meier	Terman Group	64 high school students	0.60
Bryant	Binet (Terman)	100 delinquent boys	0.38
Bryant	Binet (Terman)	100 delinquent boys	0.48

Ruch and Del Manzo (19) correlated the separate Downey traits with the Terman Group Test and the Morgan Group Test using as subjects 146 high school students. They found,

Trait	A	B	C	D	E	F
r	*0.40*	*0.10*	*0.20*	*0.16*	− 0.03	0.20
	G	H	I	J	K	L
	0.11	0.14	0.16	0.33	0.34	− 0.05

They also computed the first order partials between combinations of traits taken two at a time with intelligence constant and found that they were about the same size as these above. This means that intelligence is not a common factor accounting for inter-correlation amongst the WT traits. These writers think that there is practically no relationship between the Downey traits and general intelligence score. But total WT scores do correlate rather high with intelligence test totals. There may be many reasons for this. There probably is some connection between intelligence and this thing called temperament. But what the connection or relation is, the above figures do not show.

Downey (9, p. 278) thinks that "a temperamental pattern carries with it implications concerning the *quality* of intelligence, although not indicating the *level* of intelligence." In support of this she presents four typical profiles representing three qualities or types of intelligence: (1) The slow, cautious, deliberate, accurate type; (2) the quick, impulsive, rapid-fire explosive type; (3) an irregular, temperamental type (4) a type with a low score all the way through. She presents statistics on increased scores on Alpha when the time limits are off to show that the slow groups improve the most. Then she presents her usual case studies.

Her argument is good but her evidence is not convincing. It would be very fortunate indeed to discover a device that would measure intelligence horizontally, a thing that has been talked of a great deal. It would also be an achievement to be able to measure the social and mechanical types of intelligence. But do the WT tests accomplish these things? Proof that they do has not been presented. The only thing we can say is that they test something that does not appear to be intelligence but which correlates highly with scores on intelligence tests.

5. What uses may be made of the WT tests? Downey (9, p. 299 ff) lists a number of things that may be studied with her tests. Some of them are, racial differences in temperament, medical diagnosis, court work, the feeble-minded, social delinquents, disciplinary problems, social adaptations, salesmanship ability, executive ability, social leadership, literary temperament, research ability and ability to judge character.

Some of these have already been studied with the WT tests and the general results are reported in this paper. About all that can be said is that their value as instruments of investigation remains to be determined.

6. What value have these tests as a means of character study? This question cannot be answered without stating first the particular aspect of character study one has in mind. There are obviously many different angles from which character may be viewed and many points from which its study may be approached. Before considering some of these let us see what the author of the tests thinks of them as character tests.

Downey (9, p. 60–61) says: "The will-temperament test, since it taps chiefly the innate tendencies, is only in part to be conceived as a character test." Then she goes on to say that "character is not, then, a sum-total of native tendencies; it is a product built upon these tendencies; it is the *direction* in which the native tendencies are turned. It follows that character is not subject to measurement through tests precisely the same way as native tendencies are, although one may take a cross-section of it at any particular moment. The will-temperament determines the form assumed by character, although it does not determine its content."

Unfortunately she lets the matter rest here. Her statements are not very clear and the distinction between the form and content of character is ambigious. Anyway she does not consider the WT tests as measures of character, except in part and which part she does not state.

One of the most common methods of character study is that of analysis into traits. Can character traits (assuming that there are such things) be studied profitably by the WT tests? Perhaps they can, but the above results seem to show clearly that the Downey tests do not measure any easily identifiable traits. It is very doubtful if these handwriting exercises will correlate highly with anything that could be regarded as a character trait. Whatever else they measure, they do not measure traits.

Character is commonly regarded as conduct. Do the WT tests measure conduct? About the only data we have on this point are those of Clark and he found a negative correlation between WT score and the conduct-response score of delinquent boys. But he also found a slight positive correlation (0.29) between change in conduct-response and WT score. Conduct is a complex social affair and one would hardly expect to find it correlated with simple handwriting exercises.

Character may also be regarded as the predictability of behavior. Can behavior be predicted by these tests? The results seem to show that academic success is in some instances better predicted by a combination of WT tests and intelligence tests than by intelligence alone. But when we consider the rather high correlation between some types of intelligence tests and the WT tests we wonder how much of the prophecy is due to intelligence and how much to temperament.

While it is true that the WT tests will not foretell what any person will do in a given situation, yet the general nature of his reactions may be predicted. For example, Downey would say that the will-profile will foretell whether or not an individual's responses will be strong or weak, deliberate or impulsive, aggressive or its opposite, and so on. This type of prediction is very desirable and it seems that the WT tests have definite value at this point.

Character may also be approached from the side of social attitudes. No one to the knowledge of the writer has attempted to correlate the WT tests with such attitudes as loyalty, good-will, or used it to distinguish between social and anti-social interests. It may very well be that persons with anti-social impulses will exhibit different profiles from those with social impulses. The nearest approach we have to this is the work done with delinquent boys at the Whittier State School. No differentiating profiles were obtained, but there were striking differences between the total scores of delinquents and non-delinquents.

Perhaps the greatest value of the WT tests lies not in their usefulness as measures of character or anything else, but rather as research methods. Temperament is bound to play an important part in all character studies. Moreover, we cannot hope to make much progress in this direction until we have secured some reliable methods of measuring human energies and the ways in which they are expressed.

REFERENCES

(1) Bryant, E. K.: The "Will-Profile" of delinquent boys. *Journal of Delinquency*, vol. 6, 1921, pp. 294–309.

(2) Bryant, E. K.: Delinquents and non-delinquents on the Will-Temperament Test. *Journal of Delinquency*, vol. 8, 1923, pp. 46–63.

(3) Clark, W. W.: Supervised conduct-response of delinquent boys. *Journal of Delinquency*, vol. 6, 1921, pp. 387–401.

(4) Downey, June E.: The Will-Profile. Department of Psychology Bull. No. 3, Univ. Wyoming, 1919.

Dr. Marie McGrath. By standardizing her tests "after the order of the intelligence tests" she hoped to be able to measure kinds of moral knowledge and possible mental deficiencies in this field.[4] She employs "a series of tests and exercises, a group of pictures, and a number of little stores."[5] No attempt was made to correlate intelligence or conduct with an individual's score. Dr. McGrath gives percentages of successful answers which presumably represent the performance of an "ordinary" child at each age.

All the above tests have implied the supposition that moral behavior and moral judgments are closely related. If this were true, we would be justified in assuming that such tests afford a trustworthy indication of the subject's character. One method of critically evaluating such an hypothesis would be to administer the same test to two groups one delinquent and the other non-delinquent, and to compare the results. The Judge Baker Foundation offers an opportunity for carrying out such an investigation. Here, extensive studies are made, not only of young people presenting actual conduct problems, but also of those brought for vocational guidance. The latter constitute an excellent check group in the attempt to ascertain how adequately verbal responses give a clue to the individual's moral calibre.

For this investigation we have used the scores obtained by approximately 2000 children on the Terman Fables test. The ages range from eleven to eighteen years. This test was given at the Judge Baker Foundation as a part of the Stanford-Binet scale. These little stories of concrete situations involving moral values seem analogous to other tests of ethical discrimination. Moreover, Terman in his revision of the Binet-Simon scale, states that this test is particularly valuable in the mental examination of delinquents.[6]

---

[4] McGrath, M. C. A Study of the Moral Development of Children, *Psychological Monograph*, vol. XXXII, no. 2, 1923, p. 3.

[5] *Ibid.*, p. 11.

[6] Terman, L. M. Measurement of Intelligence. Houghton Mifflin Co., p. 299.

According to his norms the fables test is passed at twelve years if the child's score is four; at sixteen years if his score is eight. Since no provision is made by Terman for children between these ages, we evolved for the purpose of this study an arbitrary scale of values. A score of 5 was counted successful for thirteen years, 6 for fourteen years, and 7 for fifteen years. This group of thirteen, fourteen, and fifteen year-olds will hereafter be referred to as the doubtful group. The records of 1443 delinquents and 455 non-delinquents were used.[7] Following the Terman classification of I.Q.'s obtained on the Stanford-Binet, each subject was placed in one of the following groups: normal (i.e., I.Q. 90 or above), dull, (i.e., I.Q. 80 to 90) defective, (I.Q. below 80).

TABLE 1

	1443 DELINQUENTS	455 NON-DELINQUENTS
	per cent	*per cent*
(928) Successful	48	54
(625) Unsuccessful	35	27
(345) Doubtful	17	19

Table 1 shows the percentage of successful, unsuccessful and doubtful performances for delinquents and non-delinquents.

We find here that the differences are so small as to be considered practically negligible. Of the successful, there is a difference of only 6 per cent in favor of the non-delinquents, and of the unsuccessful, a difference of only 8 per cent in favor of the delinquents. Thus our results do not offer any basis for differentiating delinquents and non-delinquents, according to their ability to form moral judgments.

The division of the groups, according to mentality, shows much more significant results, as can be seen from table 2.

---

[7] *Delinquents,* i.e., those apprehended for such offenses as larceny, sex misconduct, petty misdemeanors, etc. *Non-delinquents;* i.e., those referred to the Judge Baker Foundation for educational or vocational advice.

It is to be noted that the difference here is not between delinquent and non-delinquent but between normal and defective children. This can readily be seen from the charts. Fifty-six per cent of the delinquents and 57 per cent of the non-delinquents failing to succeed on this test were mentally deficient. On the other hand, of those succeeding, only 3 per cent of the delinquents and 1 per cent of the non-delinquents fell into that category. Furthermore, of those succeeding, 82 per cent of the delinquents and 88 per cent of the non-delinquents graded as normal.

In the tables are also included the doubtfuls, together with the successes and failures into which their scores were resolved.

TABLE 2

Performance of delinquents and non-delinquents on fables classified according to mentality (1898 cases)

I.Q.	90 AND ABOVE		80 TO 90		80 AND BELOW	
	Delinquents	Non-delinquents	Delinquents	Non-delinquents	Delinquents	Non-delinquents
	per cent	*per cent*	*per cent*	*per cent*	*per cent*	*per cent*
Successes (928)	82	88	15	11	3	1
Failures (625)	17	17	27	26	56	57
Doubtful (Resolved) (345):						
Successes	39	75	51	18	10	7
Failures	38	42	36	30	26	28

These are given at their face value and were not taken into consideration in our conclusions. We do not believe that the recorded scores between 4 (twelve year success) and 8 (sixteen year success) are reliable. Moreover, our scale is only arbitrary and not borne out by experimental standardization.

It is interesting to note that there are 10 per cent more normals in the non-delinquent than in the delinquent group. In view of the apparent high correlation between mental normality and moral generalizations, this may account for the slightly higher percentage of successes in the non-delinquent group.

It is clear, at least in regard to failures, that the ability to recognize a moral situation is no basis for the detection of delinquent and non-delinquent trends. This is shown in the practically parallel percentages for the two groups. The differentiation is one of *mentality.* Terman implicitly and perhaps intentionally agrees with our conclusions. In illustrating the performances of delinquent children, he quotes

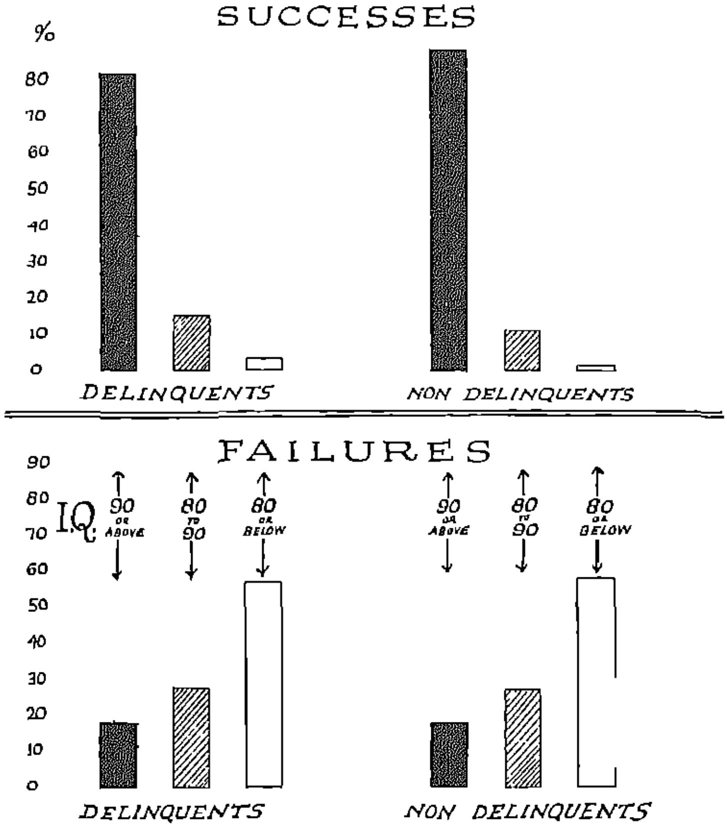


Fig. 1

only the responses of feeble-minded delinquents. We contend that feeble-minded *non*-delinquents would have ranked just as poorly.

This is well illustrated in the case of A. T., a girl of seventeen, who was referred to us for vocational advice. According to the Terman classification, she ranked as a high-grade moron (I.Q. 67). This girl was non-delinquent, and had shown, both at home and at school, good characteristics. Her score on the fables was 0. Asked what lesson we might learn from the story of *Hercules and the Wagoner*, she replied, "It teaches us to stay with the driver;" from the story of the *Milkmaid and her Plans*, all she learned was, "Teaches us not to spill our milk."

To take a less extreme example, we note the case of S. G., a sixteen year old girl who was referred to us for vocational advice and general adjustment. There was no question of delinquency. This girl was considered normal, but of poor mental ability (I.Q. 80). However, she scored but 2 on the fables, which is a failure for her age. To the story of the *Fox and the Crow* she replied laconically, "Selfishness;" from the story of the *Miller, his Son and the Donkey*, she deduced the lesson that "We should think more of others."

In direct contrast to the above cases, our files contain many examples of normal individuals with delinquent careers, who make excellent generalizations from the fables. We cite L. K., a colored boy of eleven, who had been stealing for three or four years, with bad companions. No insignificant factor in his obtaining an I.Q. of 100 was his sixteen year credit on the fables. Perhaps his most interesting response was to the story of the *Farmer and the Stork*. "You should keep away from people who steal," he replied glibly.

An analogous case is that of M. P., twelve, complained of because of excessive stealing and truancy and continued companionship with bad boys. This boy, of very good general ability (I.Q. 108), scored 6 on fables. To the fable of the *Milkmaid and her Plans*, he said, "Never think of what you're *going to have, before you get it,"* and to the *Farmer and the Stork*, "Never go with robbers, you'll get the blame."

It is needless to give further examples to illustrate this point, namely, that the ability to form moral judgments correlates very much more highly with general mental ability than with moral behavior.

From this critique of the fables, we are, of course, in no position to deny the *possibility* of a test of moral development. Our results do, however, make us suspicious of all tests having as their underlying principle the assumption that moral judgments offer a reliable estimation of moral integrity. We are convinced that verbal judgments of moral situations are an index of the individual's intellectual and social apperceptions and not his moral character.

THE FREE ASSOCIATION METHOD AS A MEASURE OF THE EFFICIENCY OF ADVERTISING

JOSEPH L. HOLMES

Columbia University

The method of controlled association has been used in the investigation of advertising by Hotchkiss and Franken[1] and by Geissler[2] with such conclusive results as to suggest the desirability of testing the free association method as a measure of the strength of association between firm name or brand and product or to use the term employed by the first of the mentioned investigators, "mental dominance." Uncontrolled association was used by Heller and Brown[3] in studying car cards but their method was far more cumbersome and much more limited in its applicability than the controlled association method or the method about to be described.

The present experiment was conducted as follows. One hundred men and one hundred women students of Columbia University and one hundred men students of Dartmouth College and fifty men students of Harvard University were used as subjects. Each one was given a multigraphed list of twenty firm names each name being that of a manufacturer of a nationally advertised product, among them being five manufacturers of toilet articles. The following directions were read to each group;

> This is an experiment in free association. Below are twenty names. Please look at each name and immediately write in the space following

---

[1] Hotchkiss and Franken. The Leadership of Advertised Brands. New York, 1923.

[2] Geissler, L. R. Journal of Applied Psychology, 1917, vol I, pp. 46–60.

[3] Heller and Brown. University of California Publications in Psychology, 1916.

the name the word or words that come to your mind. In some cases the ideas that come to your mind can be expressed in a few words. In other cases you may want to use a sentence or two or more sentences, in which case do so. You will have all the time you need so please try to give some response to every name. If any of the names remind you of nothing leave the spaces after them blank.

TABLE 1

Irrelevant responses and failures to respond

	BRAND				
	A	B	C	D	E
Columbia women	7	7	6	15	52
Columbia men	8	14	10	16	38
Harvard	2	7	2	13	3
Dartmouth	5	10	16	14	38
Total	22	38	34	58	131
Order	1	3	2	4	5

TABLE 2

Number of different responses

	BRAND				
	A	B	C	D	E
Columbia women	23	9	12	24	8
Columbia men	17	12	17	29	5
Harvard	14	8	8	13	4
Dartmouth	16	9	11	24	5
Total	70	38	48	90	22
Order	2	4	3	1	5

The responses to the names of the producers of toilet articles were tabulated with reference to frequency giving tables similar to those of Kent and Rosanoff.[4] The fifteen other names in the list were those of manufacturers of diverse products and the responses to them are not comparable. They were used to avoid a mental set which might have been established had only names of producers of toilet articles been used as stimuli.

---

[4] Kent and Rosanoff. A Study of Association in Insanity. Amer. Jour. Insanity, 1910.

This method as compared with the controlled association method seems to have the following advantages. First, the associations tested are the identical ones in which the advertisers have been drilling the buying public, that is, they are to be reproduced in the same direction in which they have been formed.[5] Second, this method gives in addition to a measure of "mental dominance," which is gotten by a comparison of the tables, one with another, an additional measure of the degree to which each of the products of any one manufacturer is known.

Assuming that the frequency of irrelevant responses and failures to respond is a measure of "mental dominance" the standing of the various brands is indicated in table 1.

Table 2 shows the variety of responses or the number of *different* responses given to each stimulus. Omitting D which is the name of a company marketing a great variety of pharmaceutical products in addition to toilet articles it will be seen that the order of frequency is the same as that for irrelevant responses and failures to respond given in table 1. At first it might seem that the number of different responses is determined by the number of products marketed by the company. This is not necessarily entirely so. However the present data do not conclusively determine this. Before further discussion it should be stated that these data were gathered for purposes other than those being discussed in the present report, which is only intended to be suggestive of the possibilities of the method.

Tables 3, 4 and 5 show the analyses of the frequency tables in regard to three products. Table 3 shows the frequency of the response "talcum powder" and its synonyms, and the order of "mental dominance" for each brand. Tables 4 and 5

---

[5] Adams, Henry Foster. Advertising and Its Mental Laws. New York, 1921, p. 181.

TABLE 3

Frequency of response "talcum powder"

	BRAND				
	A	B*	C	D	E
Columbia women	0	0	78	10	3
Columbia men	2	1	51	5	1
Harvard	0	0	35	5	1
Dartmouth	0	0	37	5	0
Total	11	1	201	31	5
Order	3	5	1	2	4

* Does not market talcum powder.

TABLE 4

Frequency of response "shaving soap"

	BRAND				
	A	B*	C	D*	E
Columbia women	2	0	12	0	43
Columbia men	21	0	41	0	56
Harvard	18	1	20	0	46
Dartmouth	20	0	37	0	61
Total	70	1	110	0	206
Order	3	4	2	5	1

* Does not market shaving soap.

TABLE 5

Frequency of response "tooth paste"

	BRAND				
	A	B*	C†	D	E
Columbia women	43	0	2	20	1
Columbia men	33	0	2	10	0
Harvard	27	0	1	15	0
Dartmouth	60	0	3	26	0
Total	163	0	8	80	1
Order	1	5	3	2	4

* Does not market tooth paste.

† Formerly marketed tooth paste.

show the same results in regard to "shaving soap" and its synomyms and "tooth paste" and its synonyms. Here then is a measure of "mental dominance" similar to that secured by other investigators. In an experiment where a measure of "mental dominance" was aimed at specifically there would be fewer zeros than in these tables for the names of manufactures of identical products would be used as stimuli. The measure of the degree to which each of the products of any one manufacturer is known is determined by the frequency of

TABLE 6

Stimulus—brand "A"

22	Tooth paste (and its synonyms)
20	Soap
13	Toilet articles
12	Shaving soap (and its synonyms)
6	Shaving cream and tooth paste
9	Irrelevant
5	Perfumes
3	Tooth paste, tooth powder, soaps and make-up articles
2	Powder
2	Tooth paste and other toilet articles
2	Shaving cream and other toilet preparations
1	Irrelevant and shaving cream
1	Soap and shaving cream
1	Irrelevant and perfumery
1	Toilet articles and a product recently marketed by the company and having a trade name
100	

the responses. Table 6 is one of the twenty frequency tables. At the top is the stimulus word, the name of the company or brand and before each response is the number showing how many of the hundred men subjects in the group gave each response.

A comparison of the frequency tables derived from the data gotten from the three groups of men shows no significant differences. As to sex differences the women show a slight tendency to give more than one response to a name, for example instead of responding to the stimulus "Brand A" with "soap" they more often than the men respond "soap and other toilet articles." They also give a greater variety of responses and more irrelevant responses than the men. Also as might well be expected their frequency of responses is greater for the products more used by women and the same is true in regard to the men, for example, face powder and shaving cream.

The writer is grateful to Professor Starch of Harvard University and to Professor Moore of Dartmouth College for their assistance in gathering the data reported on here.

THE INFLUENCE OF FRIENDSHIP UPON PERSONAL RATINGS[1]

EUGENE SHEN

Stanford University

While Rugg[2] lays down thorough acquaintance as one of the three essential conditions for reliable rating, Knight[3] concludes from an investigation, "It is in the direction of truth to discount the ratings of judges when acquaintance has been long." Acquaintance undoubtedly affects personal rating in two opposing directions: It means opportunity for close observation, at the same time it gives one a bias.

In a study on personal ratings, twenty-eight individuals who had been class-mates for at least three years were requested to rank one another with respect to friendship in addition to eight different traits (which were: *intellectual quickness*, *intellectual profoundness*, *memory*, *impulsiveness*, *adaptability*, *persistence*, *leadership*, and *scholarship*). For reasons which do not concern us here, the twenty-eight men were divided into four groups, each rating a combination of only four of the eight traits. And because two persons failed to turn in their ratings, our final data consisted of twenty-six series of ranks in friendship and thirteen series of ranks in each of the eight traits. Except those on friendship, the ranks were converted into scores in terms of standard deviations of a unit normal distribution, and ratings by judges of the same group on the same trait were averaged. The reliability for the average ratings ranged from

---

[1] Slightly modified from parts of a thesis written under the direction of Prof. Truman L. Kelley, to whom grateful acknowledgment is due for helpful suggestions and constant encouragement.

[2] Journal of Educational Psychology, XII: 425–438, 485–501, 1921, and XIII: 30–42, 80–93, 1922.

[3] Journal of Educational Psychology, XIV: 129–142, 1923.

0.62 for *impulsiveness* to 0.91 for *scholarship*, all except *impulsiveness* having a reliability well above 0.80.

If we take the average rating of an individual by the group as a criterion, and call a deviation therefrom of any individual rating an error, we can compare the errors with reference to varying degrees of friendship. Because our data show considerable grouping in the ratings of friendship, especially toward the lower end of the ranks, to keep all the twenty-seven different degrees would clearly be unprofitable. We shall therefore divide the ranks into five groups: 1–3, 4–8, 9–15, 16–24, and 25–27. To be sure, the division is arbitrary, but it is based upon the not unreasonable assumption that one has few very intimate friends and more mediocre ones. The last small group is intended to provide for a few persons whom the judge may for any reason dislike in particular.

TABLE 1

	RANKS IN FRIENDSHIP				
	1–3	4–8	9–15	16–24	25–27
Intellectual quickness	0.38	0.45	0.39	0.38	0.41
Intellectual profoundness	0.42	0.39	0.39	0.41	0.50
Memory	0.43	0.47	0.47	0.47	0.49
Impulsiveness	0.62	0.54	0.59	0.52	0.61
Adaptability	0.53	0.57	0.53	0.56	0.73
Persistence	0.56	0.52	0.54	0.58	0.67
Leadership	0.44	0.36	0.39	0.37	0.47
Scholarship	0.42	0.30	0.40	0.41	0.51
Average	0.47	0.45	0.46	0.46	0.55

Table 1 gives the average errors of thirteen judges on each of the eight traits. It shows that with the possible exception of ranks 25–27, friendship has not affected the accuracy of the ratings in any consistent manner. Even in the case of ranks 25–27, the average errors are not larger than for the other ranks to a very significant extent. In groups like the present, then, where all members have been class-mates for at least three years, especially intimate friendship does not seem to render knowledge of one another more accurate.

We shall now turn to a consideration of the systematic errors, to see whether the judges tend to rate more intimate friends consistently higher or lower than they rate less intimate ones. Whereas in the previous case we call any deviation from the average rating an error, we shall now distinguish between plus errors as errors of overestimate and minus errors as errors of underestimate. Such errors are summarized in table 2.

Table 2 shows that with the exception of *impulsiveness*, there seems to be a consistent relation between friendship and tendency toward overestimation. Since the purpose of the study was explained and absolute confidence promised when the rating blanks were sent out, there is no reason to suppose that the tendency is due to conscious discrimination according to friendship. Rather, it is due to a genuine illusion. Our best friends occupy such a favorable position in our view that whenever we compare them with others, we are comparing, as of the fan-shaped figures of Wundt, a longer side of one with a shorter side of the other. The effect is inevitable.

TABLE 2

	RANKS IN FRIENDSHIP				
	1–3	4–8	9–15	16–24	25–27
Intellectual quickness	+0.27	+0.12	+0.03	−0.13	−0.24
Intellectual profoundness	+0.26	+0.09	+0.03	−0.10	−0.12
Memory	+0.26	+0.11	+0.09	−0.12	−0.21
Impulsiveness	−0.06	+0.09	−0.10	+0.05	−0.10
Adaptability	+0.14	+0.10	−0.09	−0.04	−0.18
Persistence	+0.19	+0.16	+0.00	−0.14	−0.21
Leadership	+0.12	+0.10	+0.02	−0.09	−0.17
Scholarship	+0.20	+0.07	−0.05	−0.03	−0.17
Average	+0.17	+0.11	−0.02	−0.08	−0.18

However, it should be noted that the systematic errors are comparatively small. Large chance errors will still remain after a regression according to the systematic errors. When we conclude that there is a definite tendency toward overestimation according to friendship, we must remember that much the larger part of the errors is due to other unknown factors.

A SUGGESTION FOR AN EXPERIMENT ON THE EFFECT OF NOTE-TAKING IN LECTURE COURSES

SADIE MYERS SHELLOW

Smith College School for Social Work

Experiments have definitely proved that memory can be improved through the recitation method. Educators had worked on this assumption long before experimental psychologists had discovered its truth by testing it objectively. In grammar schools, high schools, and many courses in the undergraduate work of colleges and universities classes are conducted on the recitation plan. But as one progresses more intensively in a field, the recitation plan gives way to the lecture course. This is especially true in colleges and universities and is becoming more and more widespread as the influence of the continental methods of education is felt in this country.

With the transfer from the recitation method to the lecture course there is a change in point of view as well as in the purpose of education. The recitation is a game whereby the instructor constantly tests the familiarity of the student with material assigned to be read. The lecturer pours forth floods of material every class hour, putting the responsibility for absorption more definitely upon the student. At frequent periods the test of assimilation of material is made by way of written examinations. In order that he may be sure to retain the material presented in lectures, the student early begins the habit of taking notes. It would be interesting to study different types of note-taking; students vary from those literalists who frantically write down each word as it falls from the lips of the speaker, to logicians who can reduce a whole lecture to outline form, and to symbolists for whom a single inventive diagram or schema may represent many complex ideas. There are also the chaotic note-takers who are never quite certain what is important and what is not, and whose note-books are a jumble of principles, phrases and single words. And finally there is a critic whose notes on the lecture are only sufficient to enable him to evaluate it in his own way as he goes along. Unfortunately there are too few of the latter, for such note-taking presupposes a mature attitude with considerable knowledge of the subject matter.

The problem of note-taking is twofold: The responsibility of the instructor is not less great than that of the student. The instructor feels that if the student is going to take notes he should at least put down the essential and important points. It is also a challenge for clearness and lucidity, together with a certain succinctness. Dictation of definitions is often supplied. Note-taking for the student is often at first a real task until each works out his own system, and develops a quasi-automatic method. When this is accomplished a cycle is formed, the ideas flow from the lecturer to the student in the form of auditory perceptions, then into the neuromusculature resulting in words in the note-book. How deeply these paths are worn during the hour is of course a variable matter according to differences in individuals. It is not infrequent, however, for the student to trust a great deal to his pen and very little to his own associative processes. A review for examination then, means re-reading the notes and memorizing whatever appears to be important. The feeling of dread for examinations and sense of relief after they are finished is often based on the fear that this superficially assimilated material will not be retained during the examination period.

But is the real purpose of lecturing to fill notebooks so that students may "cram" for examinations? Is not the function of education rather, to provide the student with material for thought, to acquaint him with known facts, philosophies and theories, in order that he may live intelligently? In graduate courses the aim is most often the presentation of facts accumulated in the field so that the student may have a foundation upon which to build further and references for which to turn for further information. Often the mere memorizing of these facts is not nearly so important as the developing of a method by which the student may know how to look for facts when they are needed.

But no matter which of these three purposes is in the mind of the lecturer, the means for accomplishing the end most quickly must be through securing maximum interest and attention since these are the "sine qua non" of memory.

Moreover to many lecturers, note-taking is distracting. One feels that the student is dividing his attention, especially is it annoying when one is conscious that his every word is being preserved. The effect on many lecturers is to heighten self-consciousness and detract from that spontaneity and vitality which helps so much toward making the lecture interesting and putting across the vital points.

With these ideas in mind the writer experimented during the summer of 1924 with a course in General Psychology given to graduate students at the Smith College School for Social Work. At the opening lecture the instructor told the class that she wished to carry out an experiment and asked for their co-operation. The plan announced was as follows: No student was to take any notes whatever during the lecture. At the end of forty-five minutes the lecturer would stop, and the class was then to outline in the remaining five minutes the main topics of the lecture.

The reactions of the students to the announcement were interesting. Some were greatly relieved, some quite panic-stricken since they had so firmly formed the habit of note-taking. But all agreed to give the experiment a fair trial. In order that the exercise be not too difficult at the beginning the lecturer wrote the main topics on the board as she went along. With each succeeding lecture fewer items were written out for the class. Often, however, principles were illustrated by diagrams, and wherever possible, visual as well as auditory material was given.

The effects of this system were twofold: The instructor felt freer and seemed to be more "in rapport" with the students than during any previous summer. She was able to observe their expressions to gauge by their wrinkled brows spots of unclearness in her explanations, and to clear up what appeared to be murky places, which would have been lost had the students been intent on putting down words rather than trying to assimilate without props. The students were conscious of active attention. At first their post-lecture notes were meagre, and one or two older students, who had been out of college for many years, reported that they were completely blocked. At the end of four weeks the mid-term examination was given. The only notes which the student had for review were those made at the end of each hour, together with whatever notes they had taken on the collateral reading. One of the questions of the examination was, "What is your reaction to the 'no note-taking' experiment? In what way has it been a help or hindrance?" Below are samples of the replies to this question:

I. At the end of the lecture we have an idea of the scheme as a whole, not of a few details which have persisted during the process of note-taking. If there are portions I have not quite understood or appreciated their connection, I have looked them up in the book after the lecture, as they seem to stand out more clearly than when taking notes, when in the desire to get facts down, those not understood are not remembered. It is most helpful to have a brief outline on the board, then examples given can also be remembered even at the first of the experiment.

II. The "no note-taking" experiment at once gave me the feeling that the object of the course was background—not detailed pigeon-holed knowledge. By not taking notes the main theme or backbone of the lecture is learned at the time with the important points in proper relation to one another and with little chance of substituting details or illustrations for primary points. However, there is a loss of attention to detail and illustration in the effort to keep a clear picture in mind of the salient facts in their logical relation to one another. This may perhaps be filled in by reading.

III. The experiment was difficult for me to adapt to at first, as a feeling of panic comes over one at the inhibiting of such a long established habit. However, I thoroughly approve of it, and believe that it accomplishes the desired results first by demanding more strict attention, not only to the words uttered, but to the idea behind the words, and second by fixing the material at once in the mind rather than the energy being expended in the effort of continuous writing. I should think it would result in the covering of more ground per hour than under the other system. The system is apt to leave the mind with an outline impression of the whole mass of material rather than with a sprinkling of details which must later be organized. Sometimes I miss not being able to jot down, when it is uttered, a few words which seem to me to be a particularly apt expression of an idea, or which for some reason I believe will help fix this idea in my mind. By the time I try to recall it later, it is snowed under. This is, of course, only a minor objection.

IV. I have no settled opinion upon the no note-taking experiment. I think it tends to focus your attention, so that it is an aid to memory. However, you have a distressing feeling of unsureness when you get out your notes in order to study for "exams"—but this may be a benefit in that it tends to make you hunt out discussion on the topics you find you have barely outlined—rather than passively soak up information as you did under the old method of formal note-taking.

V. The experiment is vigorous, but worth the effort.

VI. I have no settled opinion upon the "no note-taking" experiment. I can get things clearer when I take notes and find the "no note-taking" a hindrance.

VII. "No note-taking" during lectures is pleasant enough during the lectures themselves, but not in preparing for an examination. The experiments and not the one, two, three, four reasons stand out clearly in my mind. This is evidenced by my answer to question three. I remember the chimpanzee and the button, the rat and the maze, but not concisely stated reasons. I think it is an advantage when the course is considered as a whole. But it is a hindrance to have to fill in the lectures from reading which does not present the subject just as it has been presented in class. The result is a few clear ideas, but not in an orderly arrangement.

VIII. I find the "no note-taking" a hindrance. As yet I have not benefited from the experiment. The definite hindrance has been:

a. The difficulty of attention during class, especially at that time of day (immediately after lunch).

b. The difficulty of reviewing which a brief note now and then in outline form gives.

c. In an attempt to remember specific new technical phrases the attention becomes strained, split-up, and fixed on details rather than left free to take in the general ideas.

Of the thirty-five students who answered this question, twenty-two definitely declared themselves to have been more benefited than hindered (samples I, II, III, V), while in ten the replies did not indicate whether the student had been more helped or hindered (as in sample IV), while three felt that they had been definitely hindered. These three are quoted in full VI, VII, VIII).

From the analysis of results obtained it would seem that a carefully planned experiment in the field of note-taking would prove of value. The class should be divided into two groups, one allowed to take notes in the ordinary way, as a control group, and the other asked to refrain from taking notes, but to sum up the lecture at the end of each hour. The students should be paired so that the groups might be as equal as possible in respect to such factors as intelligence and grade of work. It would then be possible to make comparison of grades received in the examination, as well as an analysis of introspective reports. When the course had been half completed (providing it were a full term course, and not a summer course), the two groups could be interchanged, so that those who had not taken notes before would now take notes, and vice versa. Thus comparison of the two methods, (*a*) within groups, (*b*) on the same individual, could be made.

It is of course true that all subjects would not lend themselves so readily to this experiment. Those courses in which a large amount of definite facts must be learned, as for example, anatomy or medicine, would not be suitable for such an experiment. In order to get the best results the following conditions must be met:

1. The lecture itself must be developed in a structural way, with each point clearly brought out and well illustrated.
2. Unfamiliar and technical terms should be written on the board.
3. Wherever possible, diagrams should be used to assist the memory.
4. The amount of time necessary at the end of each hour for note-taking should be determined experimentally.
5. Collateral reading on material covered by the lectures should be assigned.

The student should be urged to:

1. Outline the lecture as fully as possible.
2. Ask questions about any points not clearly brought out.
3. Look up any material which may appear vague.

THEORIES OF SELLING[1]

EDWARD K. STRONG, Jr.

Stanford University

There are undoubtedly many theories more or less active in the minds of sellers as to how to influence others to buy. But a perusal of the literature on the subject leads to the conclusion that there are only three general theories which are definitely formulated.

The first theory is expressed by the five words: "*Attention, Interest, Desire, Action, Satisfaction.*" There is implied in this theory that the prospect must successively experience the *conscious states* of attention, interest, desire, action and satisfaction. And the emphasis is put upon establishing these states of consciousness.

The second theory comes to the selling world from behavioristic psychology. All behavior according to this view can be expressed by the formula: "*Situation-response.*" The word "situation" comprises factors external to the individual and also factors within him. Thus, the remark, "Let's have a smoke," would be an external factor and the presence or absence of a desire to smoke would be an internal factor leading to the response of smoking or not.

The expression "*Appeals-Response*" has appeared frequently in the advertising world as a substitute for "Situation-Response." When either of these formulae is employed undue emphasis always seems to be laid upon the external factors; and the emphasis has accordingly been put upon determining which appeals will stimulate.

The third theory views man as a dynamic being. Whether man buys or not depends very largely upon the internal factors within him, summed up in the word "*wants*." Unless he wants he will remain unaware, or at least uninterested, in what the seller has to present.

---

[1] Read in part before the Western Psychological Association, August 8, 1924.

Psychologists may not experience any difficulty in keeping both external and internal factors in mind when they use the expression "situation-response." But the layman without a broad grasp of the subject must apparently overlook one or the other: He can not keep both before him. As he naturally is interested himself in what he is presenting, namely, appeals to the prospect, it seems best to emphasize the internal factors in any formulation that is given to him as a guide in his thinking.

The words: "*Want-Solution-Action-Satisfaction*" express this third theory of selling. The emphasis is put upon the wants of the prospect and the seller's function is to guide the thinking of the prospect to the desired end of buying.

Consideration of each of these theories in more detail will help to make the above points clear.

SELLING EXPLAINED IN TERMS OF MENTAL STATES

E. St. Elmo Lewis formulated the slogan, "Attract attention, maintain interest, create desire," in 1898. Later he added the fourth term "get action." These four words represent four states of consciousness which must pass through the mind of the prospect before he will buy. In other words, if the prospect experiences attention, interest, desire, he will be more likely to act; and consequently an advertisement or sales talk must be planned to arouse these states in him.

Sheldon[2] is apparently responsible for adding the fifth term, "secure satisfaction." In his writings of 1905 and 1910 the term does not occur. In his correspondence course of 1911 he lists the "four states that occur in the mind of the customer in every sale" and then adds, "while these four mental states are sufficient for a sale, two more must be added for the purpose of business-building, namely, Confidence and Satisfaction." In his *Art of Selling* published the same year he lists satisfaction with the other four as "the successive steps in the process of a selling transaction."

---

[2] A. F. Sheldon, The Science of Business Building, 1905 and 1910; Salesmanship, 1911; The Art of Selling, 1911.

This slogan of Lewis and Sheldon has had a very profound effect upon the selling world. At the time it appeared there was no clear and definite formulation of what was implied in selling. Books on selling and advertising were hardly more than a jumbled collection of opinions on a great variety of topics. The formula has caused order to come out of chaos.

The majority of articles and books written since 1907 have endorsed the slogan in one form or another. And many speeches before advertising men and salesmen clearly show its influence. The writer would estimate that 90 per cent of all those engaged in these fields are consciously or unconsciously influenced by this point of view.

Several writers have more or less stressed mental states without subscribing to the above five-step formula. Scott[3] writing in 1903 and 1908 believed that all the mental activities of man were involved in the process of buying, and discussed advertising under the chapter headings of: Attention, association of ideas, suggestion, fusion, perception, apperception, illusion, imagery, memory, feeling, instinct, will, habit, and the like. But Scott made no attempt to formulate a simple expression of the selling process. The same is true of another psychologist, Adams,[4] who wrote in 1916. Most of the material in his book appears under the chapter headings of sensation, attention, association, fusion, memory, appearance (feeling) and action. Adams states in his preface that "the behavioristic standpoint has been adhered to throughout" and any behaviorist could interpret his material from that standpoint. But the ordinary reader gets the point of view from his work, as well as from that of Scott, that mental states must be emphasized in selling. Eastman[5] organizes his material in Part I, making the sale, under the following heads: (1) pre-approach (*a*) investigation and (*b*) preparation; (2) approach and securing attention; (3) solicitation, (4) objections and (5) closing. Parts II and III are then devoted to the processes of thinking, feeling and action. Hess[6] discusses a formula in his 1915 work but devotes little space to it. Instead he presents chapters after the order of the above on: sense experience; instincts; imagination; memory; attention; desire; habit and inhibition; and getting the will of the crowd. In his 1923 work there is no reference to the formula or to the above psychological topics. Most of the book is devoted to the personality of the salesman and the buyer.

---

[3] W. D. Scott, Theory of Advertising, 1903; Psychology of Advertising, 1908.

[4] H. F. Adams, Advertising and Its Mental Laws, 1916.

[5] G. R. Eastman, Psychology of Salesmanship, 1916.

Substitution of "Conviction" for "Desire"

Many writers have emphasized that an intellectual rather than emotional state of mind is the key to action and have substituted words like "conviction," "confidence" or "judgment" for "desire" in the slogan of Lewis and Sheldon. Thus The Little Schoolmaster[7] wrote in 1910, "Most of those who have attempted to outline the proper arrangement of a complete advertisement—that is, an advertisement that embodies all the functions—seem to follow about this order: attract attention, develop interest, produce conviction, induce action." The International Correspondence School Text on Advertising substitutes conviction and confidence for desire. But in their text on Selling[8] there appears this formulation: (1) preparation for approach, (2) approach (with no emphasis on getting attention), (3) building up the interest, e.g., "demonstration of the practicability of the proposition, making the customer understand why he should buy," (4) closing, (5) maintaining the sale and holding the customer.

---

[6] H. W. Hess, Production Advertising, 1915, Creative Salesmanship, 1923.

[7] Printers' Ink, December 1, 1910, p. 74.

[8] International Correspondence Schools, The Conducting of Sales, 1911.

Hall,[9] 1915, substitutes confidence and conviction for desire. In 1924 when addressing salesmanagers instead of advertising men he presents: attention, interest, demonstrate, clear up details and doubts, action. Osborn[10] uses the word "judgment" for that of "desire," giving the four-word slogan: attention, interest, judgment, action.

Kitson[11] keeps the word "desire" but inserts between it and "action" the word "confidence." His entire book is based directly on this six-word slogan. Ramsay[12] adds "caution" in a similar way giving the formula: "attention, interest, desire, caution, action." But unlike Kitson, he almost completely ignores it throughout his book.

Still another variation is: "attracting attention, creating desire, removing inhibitions, inspiring confidence, impelling to action."[13]

"Satisfaction" is not always included in the formula

It is not surprising that in the years immediately following Sheldon's addition of the word "satisfaction" to Lewis' formula, the word satisfaction should be omitted because at that time the objective of selling was still viewed by nearly all as merely to secure a sale. But as the years have rolled by more and more sellers have come to see that the objective of selling is not a single sale but a customer. The word satisfaction should consequently be included in any formula in order to emphasize this new objective. But many authorities have not seen this point, apparently.

The George Batter Agency is credited with the slogan: "To be seen, read and believed." Hall in 1915 added "to be remembered" and gives this four-term slogan again in his handbook in 1921. Starch[14] makes considerable use of the slogan in this form, expanded by him to: "Must be seen, read, believed, remembered, acted upon." Another expression in similar form is credited to H. C. Sheppard,[15] namely, "Make 'em look, like it, learn, and then land 'em."

---

[9] S. R. Hall, Writing an Advertisement, 1915; The Advertising Handbook, 1021; Handbook of Sales Management, 1924.

[10] A. F. Osborn, A Short Course in Advertising, 1922.

[11] H. D. Kitson, The Mind of the Buyer, 1921.

[12] R. E. Ramsay, Effective Direct Advertising, 1921.

[13] "The Five Steps of a Sale," Published by the West Coast Life Insurance Co., 1920.

Although Neystrom,[16] Whitehead,[17] Sloan and Mooney,[18] Blanchard,[19] Ruxton,[20] and Herrold[21] recognize the value of good will, they do not include the word satisfaction in their formula and present only the four-word theory—attention, interest, desire and action. Hess,[22] 1915, applies the formula to only those customers who have no desire for the goods. He adds the word "possession" after "action," but not the word "satisfaction."

An interesting deviation from the formula appears in *Knack of Selling*.[23] The salesman is told: (1) to gain attention, (2) to hold interest, (2) to keep command over the trend of the interview and (4) to keep the canvass going that makes the prospect want your goods, arouses the motive that makes him willing to buy them and closes the order.

Expansion of the formula

The slogan is carried to its fullest extent in the *Selling Process* by Hawkins,[24] Sales-manager of the General Motors. A very brief outline of the book follows:

---

[14] D. Starch, Principles of Advertising, 1923.

[15] O. C. Sheppard, He Capitalizes the Personality of His Product, Printers' Ink Monthly, July, 1921, p. 49.

[16] P. H. Neystrom, Retail Selling and Store Management, 1915.

[17] H. Whitehead, Principles of Salesmanship, 1917.

[18] C. A. Sloan and J. D. Mooney, Advertising the Technical Product, 1920.

[19] F. L. Blanchard, The Essentials of Advertising, 1921.

[20] R. Ruxton, Printed Salesmanship, 1922.

[21] L. D. Herrold, Advertising for the Retailer, 1923.

[22] H. W. Hess, *op. cit.*

[23] H. Watson, Knack of Selling, Book II, 1913.

[24] N. Hawkins, The Selling Process, 1918.

- *a*. Preparatory steps:
    1. Preparation, Chapter IV
    2. Prospecting, Chapter V
    3. Approach and audience, Chapter VI
- *b*. Presentations steps:
    1. Sizing up the buyer, Chapter VII
    2. Gaining attention and awakening interest, Chapter VIII
- *c*. Convincing steps:
    1. Persuading and creating desire, Chapter IX
    2. Handling objections, Chapter X
- *d*. Closing steps:
    1. Securing decision and obtaining signature, Chapter XI
    2. Get-away and lead to future orders, Chapter XII

Hawkins recognizes three stages of attention: (1) compulsory attention, (2) curiosity to some degree and (3) intentional or spontaneous attention; and three stages of interest (1) attentive interest, (2) associating interest, and (3) personal interest. He insists that the prospect must be taken through all these six stages before he reaches the state of desire. Regardless of whether this is true or false the writer is certain that no salesman ever can identify these six stages as they are experienced by a prospect; and no salesman ever does or can say to himself, "Now he is in stage three of attention; I must get him into stage one of interest."

Although Hawkins' book is based on this formula the writer is convinced from studying it, especially Chapter IX, that Hawkins' point of view belongs under the third theory where wants are emphasized, not states of mind. But he has clearly attempted to fit his views into this formula of attention, interest, desire, action, satisfaction.

Emphasis upon this formula decreasing

Judging from a perusal of recent literature interest in this theory is dying out to some extent. Several books like those of Sloan and Mooney on *Advertising the Technical Product*, 1920; Blanchard, The *Essentials of Advertising*, 1921; Ramsy, *Effective Direct Advertising*, 1921; Ruxton, *Printed Salesmanship*, 1922; and Herrold, *Advertising for the Retailer*, 1923; devote only a few pages to the subject. These writers outline what is to be done and how to do it. Whatever theory they have is practically excluded from their works.

Nearly all authorities still emphasize "attention" and "interest"

Although many authorities do not include the work "satisfaction" in their formulation and some writers employ "conviction" for "desire," yet all insist that "attention" and "interest" must be present and must come first.

It is true that attention and interest must be present whenever a prospect does buy. But it ought to be equally apparent that these mental states are not the goal that the seller desires to reach. Instead, the seller should aim to establish certain ideas and feelings relative to what the buyer can get out of his goods. The mistaken emphasis is responsible for many a poor advertisement and for many a poorly introduced sales presentation.

SELLING EXPLAINED IN TERMS OF APPEAL AND RESPONSE

Hollingworth[25] gives his book the sub-title, "Principles of Appeal and Response," indicative of his behavioristic point of view. Chapters I, XIII, XIV, and XV discuss the various appeals that will produce a buying response. Much space is devoted to the question as to which appeals are the strongest. The remainder of the book considers four tasks of the advertiser or seller: namely, catching the attention, holding the attention, fixing the impression and provoking the response. The theory is apparently that if the proper appeal is presented and the prospect gives it his attention, the desired response will result, especially if the process is skilfully handled.

Hollingworth makes clear to the discerning reader that the strength of an appeal is dependent upon the prospect's own instinctive and habitual ways of acting. In other words, appeals are good or poor depending upon the internal factors affecting the individual at the time. But the ordinary reader does not get this subtle point.

---

[25] H. L. Hollingworth, Advertising and Selling, 1913.

Parts II, III and IV of Starch's book[26] are concerned with a study of the prospect, the appeals that can be made to him, and *the presentation of those appeals. His whole work stresses the* necessity of determining what appeals will be most effective in selling a particular commodity to a particular class of prospects.

So long as emphasis is put upon appeals—external factors presented by the salesman—just so long there remains the problem of making the prospect pay attention and become interested. This is true because the emphasis remains upon causing the prospect to become interested in what the salesman has to say instead of the seller adapting himself to the prospect and presenting to the buyer what he desires.

It is most natural then that both Hollingworth and Starch should continue to emphasize the getting of attention and interest. It is difficult, in fact, to get away from this point of view. The present writer tried very hard to do so but with imperfect success in his book on *Psychology of Selling Life Insurance.*[27]

The discussion in this book is presented under two main heads: The strategy and the tactics of selling. Strategy covers the preapproach or preparation for selling; tactics, the execution of the sales plan when face to face with the prospect.

The strategy is viewed as involving five steps. First of all, the prospect must be analyzed in order to discover his interests, *desires and needs. Second, the appropriate insurance to fit* his case must be decided upon. Third, the convictions that the prospect must have before he will buy must be determined, and fourth the impulses (wants) that he must be conscious of before he will buy must be determined. Fifth, the salesman must decide what appeals he must make to the prospect so that he will establish the necessary convictions and arouse the necessary impulses and so cause the prospect to buy.

The view held here is that if a given prospect becomes finally possessed of certain convictions and certain impulses he will respond by buying. Furthermore, to get these convictions and wants into the prospect's mind the salesman must present (1) ideas to establish the convictions and (2) incitements (words arousing desire) to arouse the wants. What the salesman presents are appeals or stimuli and all the wants and convictions that come to mind within the prospect are motives for buying, tending to cause the response.

---

[26] D. Starch, *op. cit.*, 1923.

[27] E. K. Strong, Jr., Psychology of Selling Life Insurance, 1923.

There is a frank recognition in this theory that the prospect will not buy until he is possessed of certain ideas and wants; that these must be guessed at by the salesman and that then he must cause them to possess the mind of the prospect. Although the writer is still convinced that this is sound psychology he is now certain that the theory as outlined above is too complicated for the *average* salesman to comprehend or use.

SELLING EXPLAINED IN TERMS OF WANT AND SOLUTION

In the section of the work of Tipper, Hollingworth, Hotchkiss and Parsons[28] devoted to a discussion of the theory of selling appear these five main heads:

1. Fundamental needs of man.
2. The commodity.
3. The establishment of association between commodity and need—the creation of mental habits of such a sort that the feeling of the need at once suggests to the mind of the individual the commodity in question.
4. The making of the association dynamic.
5. The securing of vividness of impression.

The emphasis here is upon what the prospect wants or needs and the aim of advertising is to associate this need with the seller's goods. The starting point is what is in the mind of the prospect, not appeals to be put into that mind. This sounds a new note in selling literature. But for some reason little attention has been given it.

Charters[29] pictures the prospect having a need to be satisfied or a problem to be solved. He defines the art of retail selling as the art of helping the customer to define his need, and to select the articles which will satisfy that need. The five steps of purchasing, according to Charters, are:

---

[28] H. Tipper, H. L. Hollingworth, G. B. Hotchkiss, F. A. Parsons, Advertising, 1915, Chap. IV.

[29] W. W. Charters, How to Sell at Retail, 1922.

a. The customer tries to get a clear idea of what he wants.

b. He looks at many articles which might possibly satisfy him.

c. He studies these in the light of his standards.

d. He selects the one which seems to meet them best.

e. He sizes up his purchase and experiences a feeling of satisfaction or of dissatisfaction.

Here we have a picture of selling where the prospect is very active because he has a want to be satisfied and the salesman's job is to aid the prospect in his endeavors to solve his problem. The emphasis is quite different from that expressed in the other two theories of selling.

In Herrold's[30] text are to be found passages supporting all three of the theories discussed here. No one can read his work, however, without appreciating his emphasis upon the buyer's wants. Thus, under the heading "Adopt the customer's viewpoint" he says:

> It follows, then that in determining what to say about the goods, merchandise should be looked at from the buyer's viewpoint: that in advertising, as in selling across the counter, it is necessary to show the person who is being influenced how the goods will benefit him. When selecting the selling points, think why the reader should buy the article. Study it from his standpoint. Why whould he want it? What needs of his will it satisfy? What will it do for him?

The present writer has formulated a theory of selling in terms of Want, Solution, Action and Satisfaction.[31] Selling with such a formula in mind becomes a process of calling to the prospect's mind some one or more of his wants and then showing how they may be satisfied by buying the seller's goods. The term "action" implies that the seller must more or less force the issue at the end; also that the prospect must not only buy but use the commodity, and that the seller must make sure he can and will use it correctly. "Satisfaction" is, finally, most important, because unless the goods measure up to expectations there will be no repeat orders.

---

[30] L. D. Herrold, *op. cit.*

[31] E. K. Strong, Jr., Psychology of Selling and Advertising, in press.

When "want" is emphasized first, "getting attention and interest" ceases to be a major problem

Whenever one of a prospect's wants is called to his mind, his attention and interest are automatically secured, and are essentially relevant. Consequently, the terms attention and interest can be eliminated from selling and advertising literature. In every case, the problem reduces itself to: "What wants does he have? How can I get them into his mind?"

There remain a number of minor problems that have always been discussed under the headings of getting attention and interest. They had better be considered from other angles. "How to introduce oneself," "Whether to shake hands," and all the rest belong under such a heading as "Starting the Interview." And problems of "Size of the advertisement," "Eye-movement," "Balance," "White-space" belong under such a heading as "Layout" or "How to present the proposition in the best manner."

NOTES AND NEWS

COMMEMORATIVE MEETINGS IN MEMORY OF G. STANLEY HALL AND EDMUND C. SANFORD

Last year Dr. Sanford made plans for a series of commemorative meetings in honor of the life and work of Dr. Hall and this plan included scientific conferences to be participated in especially by the alumni of Clark University actively engaged in research in genetic psychology.

On account of the death of Dr. Sanford, which occurred November 22, 1924, it seemed best to the Committee to include in this series of meetings a commemoration of Dr. Sanford as well as of Dr. Hall. The following program shows the scope of the meetings:

Chairman: Colin A. Scott, Ph.D., Professor of Education, Mt. Holyoke College.

Commemorative address for G. Stanely Hall, by the Chairman.

Commemorative address for Edmund Clark Sanford, by Mary Whiton Calkins, Ph.D., Professor of Psychology, Wellesley College.

Reading of letters from friends and former students.

Informal addresses by alumni of the University.

Founder's Day Exercises

Chairman: Wallace W. Atwood, Ph.D., President of Clark University.

Address "Changing Ideals of Education," by Frederick C. Ferry, Ph.D., President of Hamilton College.

Scientific Meetings

The Curiosity Theme in Norse Folklore, by Albert N. Gilbertson, Ph.D., Malden, Mass.

The Biography of a Nine Months Old Infant, by Edgar James Swift, Ph.D., Professor of Psychology, Washington University.

Subject to be announced, by Wolfgang Koehler, Ph.D., Professor of Psychology, University of Berlin.

Wanted: A Theory of Child Play, by George E. Johnson, A.M., Professor of Education, Harvard University.

Tactual Interpretation of Speech, by Robert H. Gault, Ph.D., National Research Council, Washington, D. C.

A Clinical Psychology of Infancy, illustrated by stereopticon, by Arnold Gesell, Ph.D., Professor of Child Hygiene, Yale University.

Some Human Factors in Mental Measurement, by James P. Porter, Ph.D., Professor of Psychology, Ohio University.

Genetic and Experimental Psychology in Relation to the Public School Clinic, by George E. Dawson, Ph.D., Professor of Psychology, Springfield College.

The Genetic Method in Present Day Psychological Research, by William F. Book, Ph.D., Professor of Psychology, Indiana University.

The special memorial meeting on Sunday afternoon consisted not only of the special addresses in honor of Dr. Hall and Dr. Sanford, but letters from alumni of the institution were read including most remarkable tributes to the inspiring personality and pioneer work of Dr. Hall and the remarkable character and work of Dr. Sanford as teacher and investigator.

While conditions prevented the presence of two or three speakers, the scientific conferences included the account of some remarkable contributions to genetic psychology, and the whole series presented an appropriate scientific memorial of the life and work of Dr. Hall.

NATIONAL RESEARCH COUNCIL

The division of Anthropology and Psychology has several very active committees:

1. That on State Archaelogical Surveys has been remarkably successful in arousing interest in quite a number of states in their local Indian remains.

2. That on Scientific Problems of Human Migration is now completing its second year of active work. In this are enlisted a large number of investigators in different parts of the country, working partly for the development of test methods that are applicable to subjects of various cultural backgrounds, partly upon race characteristics and the effect of race mixtures.

3. That on Vestibular Research is bringing together a group of investigators working at different phases of the problem of the inner ear.

4. The committee on the possibility of teaching the deaf to perceive vocal sounds through skin stimulation, and its investigator, Dr. Gault, have good hopes of making a real contribution to the welfare of deaf people.

5. Several committees have to do with problems of personnel. The latest achievement of the Committee on Personnel Research is the preparation of the new college entrance tests which have been given to many thousands of students the current year, and the returns from

which are now to be thoroughly worked over. The Committee on College Student Personnel has formulated a feasible scheme for coöperation between colleges in the research necessary before suitable vocational advice can be given by college advisors to graduating students. The Committee on the Gifted Student Problem is keeping up a steady effort to improve conditions in the colleges for the gifted student by the development of honors courses and by giving the student a perspective of the opportunities in research for a career.

6. A new committee on the Psychology of the Highway has just begun operations. It will coöperate with state and national bodies interested in the reduction of automobile accidents.

A preliminary gift of $3,000,000 for the endowment of the John Simon Guggenheim Memorial Foundation Fellowships was announced recently by Simon Guggenheim, former United States Senator from Colorado, and his wife.

In a statement explaining the purposes of the Guggenheim Fellowships, Simon Guggenheim said: "I want to supplement the great Rhodes Foundation by providing a similar opportunity for older students of proved ability, and for women as well as men. Furthermore, I want to make it possible for these persons to carry on their studies in any country in the world where they can work most profitably."

The Foundation is a memorial to the son of Senator and Mrs. Guggenheim, who died on April 26, 1922. The Foundation offers to young men and women world wide opportunities under the freest possible conditions to carry on advanced study and research in any field of knowledge, or opportunities for the development of unusual talent in any of the fine arts including music.

No age limits are prescribed. Appointees, however, must be old enough to have shown marked ability in their particular subject. It is expected that ordinarily they will not be younger than 25 or older than 35 years.

The amount of money available for each fellowship will be approximately $2500 a year, but may be more or less, depending on individual needs.

While appointments will be made ordinarily for one year, plans which involve two or three years' study will also be considered and in special cases fellowships will be granted for shorter terms with appropriate stipends.

The Joint Information Bureau in Moscow is in constant communication with most of the scientific and cultural establishments and institutions of the Soviet Union and is in a position to facilitate regular contact as well as an exchange of scientific periodical and non-periodical publications between the interested institutions, groups and individuals of the Soviet Union and the other countries. Personal exchange of their works between scientists engaged in similar research activities could also be arranged as well as the publication of unpublished manuscripts and articles in the general periodical and special press.

At the request of the Joint Information Bureau in Moscow the Russian Information Bureau in Washington has undertaken to represent the Moscow Bureau in its endeavors which will no doubt be in the mutual interest of both countries.

The interested institutions, organizations and individuals may apply to the Russian Information Bureau, 2819 Connecticut Avenue, N.W., Washington, D. C.

The Eyesight Conservation Council of America, Times Building, New York City, announces that it is prepared to furnish lantern slides and lecture material on eyesight conservation at a nominal price to those who can make good use of such material with classes of students and other groups who would profit by and be interested in such information. While the Council desires to avoid indiscriminate free distribution and use of its material, it will lend the slides without fee to those without available funds.

Dr. A. J. Snow, of Northwestern University, is now devoting most of his time to consultation work in connection with the selection of taxi-cab drivers, with offices at the Yellow Cab Company, 57 East Twenty-first Street, Chicago.

Clark University recently announced the election of Dr. Walter S. Hunter to the G. Stanley Hall Chair of Genetic Psychology. Dr. Hunter, formerly Professor of Psychology at the University of Kansas, will begin his work at Clark at the opening of the next academic year.

BOOK REVIEWS

M. V. O'Shea. *The Child: His Nature, and His Needs.* The Children's Foundation, Valparaiso, Ind. Pp. 516.

This contribution of the Children's Foundation is a volume of 500 pages of joy for the one interested in child welfare and social progress. It is "A Survey of the Present Day Knowledge Concerning Child Nature and the Promotion of the Well-being and Education of the Young." The volume indicates an awakening in child education. Men and women have been selected to contribute to this volume who are familiar with the problems of the parent, the teacher, and the social worker. In the problem assigned them they have had full freedom to express themselves as they wished. Each chapter is very brief and to the point.

In Part I, which is a discussion of the present status of our knowledge of child nature, Dr. Baldwin gives a good view of the way in which the gap between our knowledge of child nature and the training of school children may be bridged. Miss Whitley of Columbia University gives us a clear picture of the child's instincts and impulses. She says that the child needs direction and not suppression; that the growth of children is a significant study for the educator, and that we must keep in mind the development of the child's motor control. Dr. Dearborn's chapter on the development of intellect in childhood and youth deals chiefly with the intellectual development as a whole in connection with the individual organism from the physical standpoint. The chapter on the child's moral development written by Mr. Neumann is of great value to the social worker who wants to gain a real insight into the impulses of the individual. Mr. Bolton has given a valuable picture of the social traits of childhood and youth. The teacher will find much help in the chapter written by Dr. Kirkpatrick. He includes in his study language, drawing, and music as the child's mastery of the arts of expression.

Part II deals with the present status of our knowledge of the child's well-being. Dr. Goddard discusses the need of bridging the gap between our knowledge of the child's well-being and the care of the young. Every chapter in part II is valuable to the educator interested in the care and treatment of exceptional children. Dr. Goddard says, "There is much yet to learn but we already know enough, if we put it into practice, to transform civilization within a generation or two." Dr. Emerson shows a wonderful insight into the need of scientific control of nutrition. He says that improved nutrition brings increased power of mind into operation. Mr. White of St. Elizabeth's Hospital, Washington, discusses nervous and mental hygiene. He says that the child is not a small adult but is a different being and must be treated accordingly; and also that children are not alike but different from one another. To the nurse as well as to the teacher the chapter on sense defects by Professor Winslow of the Yale Medical School is very valuable. The subject of delinquency is discussed by Dr. Healy. He says that in the treatment of delinquency both the child and the environment must be considered. Dr. Arnold Gesell discusses the care of intellectually inferior children, and Dr. Leta S. Hollingworth, intellectually superior children. Dr. Winfield Scott Hall discusses the hygiene of the adolescent. He offers a suggestion which would be valuable to the educator of the child of that age.

Part III includes the study of the present status of our knowledge of education. Hon. J. J. Tigert, United States Commissioner of Education, discusses the need of bridging the gap between our knowledge of education and our educational practice. Dr. O'Shea contributes five chapters to this large volume of study. They are: Changing Objectives in American Schools, Changing Courses of Study, Changing Methods of Teaching and Management, Promoting the Health and Physical Development of School Children, and Extending Educational Facilities, Opportunities and Requirements.

We owe much to the trustees of the Children's Foundation for providing funds which have made this contribution possible.

M. LaVinia Warner, .
Ohio University

Second International Congress of Eugenics, Vol. I, *Eugenics, Genetics and the Family;* Vol. II, *Eugenics in Race and State*. The Williams & Wilkins Company, Baltimore, Md. Vol. I, 438 pp., Vol. II, 472 pp.

The two volumes containing the addresses and scientific papers given at the Second International Congress of Eugenics compose an authoritative study of present knowledge and theories concerning the important fields of eugenics and genetics. The papers are individually analytical in treatment and informative in context. As reference books they are invaluable; as a source of general information and interesting reading in the important field of race character and betterment, they could not well be surpassed.

The Congress was truly international in character, and in his address of welcome Henry Fairfield Osborne says, "I doubt if there has ever been a moment in the world's history when an international conference on race character and betterment has been more important than the present."

The volumes contain papers of both a genetical and eugenical nature. Genetics is a pure science dealing with the laws of heredity; eugenics is an application of the laws of genetics to human mating for the benefit and improvement of the race.

Volume I contains papers dealing with the relation of eugenics and genetics to the family. They approach the question from two viewpoints; first, the results of animal experimentation, and secondly the statistical evidence from family and social groups.

The work of the geneticist is experimental and scientific; in eugenics conclusions are drawn chiefly from statistical evidence since it is difficult to experiment in the field of human genetics. Pages 29 to 242 of Volume I contain papers giving the results and conclusions of the best-known investigators in genetics and range from studies of inheritance in unicellular organisms to inheritance of mental diseases and disorders. The complete knowledge and beliefs of current thinkers in this important field are found here.

The latter part of this volume contains reports of biological researches on families, an application of the principles of genetics to human mating. Particular families are discussed, showing the effects of inbreeding, of mating with adverse types; of disease, of differential fecundity; of attempts at control of mating. Such subjects as "Is Inbreeding Injurious," "The Effect on the Germ Plasm of Isolation in a Mountain Section," "The Oneida Community Experiment in Stirpiculture," are typical of this section of the book.

As a whole the treatment is scientific and thorough but written so as to be easily intelligible to the popular mind. Scarcely greater wealth of interesting fact and authoritative theory could be compiled in one volume.

Twenty-four plates of illustrations, tables and drawings lend interest and clearness to the volume.

Volume II contains papers showing the value of eugenics to the race and state. The world's foremost Eugenicists and Anthropologists contribute to this volume. This volume alone would have justified the appellation, "an International Scientific Race Congress."

Physical differentiation between races is emphasized in the beginning. Particular races are described and comparisons made. Relationship between weight and stature is shown among other things.

"Harmonic and Disharmonic Race Crossings," by Dr. Jon Alfred Myxen, and "Hybridization and Behavior," by E. M. Vicari, are typical of the many interesting papers in this section.

Major Leonard Darwin, English Eugenicist, then introduces the next section by a stirring address, "The Field of Eugenics Reform."

Population as a factor affecting the race is then ably discussed by expert investigators of the question.

Selective elimination, control of parent-hood, and sterilization are authoritatively treated by specialists in the field.

The relation of eugenics to disease, to education, to inheritance, and to society is critically analyzed by the foremost authorities. The reader has thus at his command the facts of genetical research combined with statistical evidence and specific examples in the field of Eugenics.

Twenty-four plates of illustrations in this volume add to the interest. Combined, these two volumes are indispensable in a reference library and are extremely interesting and profitable for general reading.

EMMETT ROWLES,
Ohio University.

LUCIEN LEVY-BRUHL. *Primitive Mentality.* New York, The Macmillan Co., London, George Allen and Unwin, Ltd., 1923. 458 pp.

This volume may be looked upon as an application of psychological principles to the study of ethnographic data. Professor Levy-Bruhl attempts an investigation of the character of the primitive mind as it is revealed and operates in the various activities of primitive peoples from practically all locations on earth. As a basis for his studies the author proceeds upon the excellent assumption that before studying the primitive mind it cannot be defined as a rudimentary form of our own; so that one may conclude at once that it is childish and almost pathological. Instead, he assumes that the primitive mind is normal under the conditions in which it is employed and further more that it is both complex and developed.

Notwithstanding this laudable hypothesis with which Professor Levy-Bruhl approaches his task the results reached by our author from the psychological study of primitive action is that there are decided differences of a fundamental sort between primitive and civilized mentality. Briefly, the primitive mind is essentially mystic and prelogical. Primitive man is not interested in natural causes but is entirely occupied with a world in which occult powers prevail.

That primitive man is much occupied with the occult and the mystic is in no wise to be denied, but that this means that primitive mentality represents a unique form of psychological functioning entirely different from our own is an interpretation that has never been validated. To believe that primitive mentality is fundamentally different from our own is to overlook the myriads of instances of mystic and prelogical mental action performed by persons in all complex civilizations. Not even the most perfect of our sciences are immune to occult and mystic thought.

J. R. KANTOR,
Indiana University.

CLARK L. HULL. *The Influence of Tobacco Smoking on Mental and Motor Efficiency.* Psychological Monograph, No. 150. Psychological Review Company. Pp. 161.

In the first chapter of this study the author presents, though briefly, a summary of previous outstanding investigations of the effects of tobacco smoking. He selected the pipe as the medium of experimentation. The work is predicated upon an ingenious control device, a thing that has presented great difficulty to other investigators. The method of procedure is empirical throughout. Smokers have been distinguished from non-smokers and interesting parallels have been drawn between the two types of performance.

The study is effected by the use of tests designed to show how smoking affects, 1. neuro-muscular processes, 2. sensory motor processes and 3. higher mental processes. For each group a series of twelve tests is given. The series is repeated four times daily including the "normal for the day." A special program for alternation of smoking and control performance was devised. In three of the twelve forms of behavior investigated, the positive effects of smoking are evident. Two of these are physiological processes (pulse rate and muscular tremor), and the third (addition), a psychological process. Results of other tests show a fair degree of reliability in pointing to tobacco effects and are entitled to consideration.

The work will, no doubt, be followed by many others before the psychological effects of smoking are fully established. The author has proved himself to be both an accomplished statistician and an able investigator. He has dealt with involved data and has succeeded in giving them a lucid presentation. His plant for conducting the inquiry was carefully devised. The data have been fully and capably handled and conclusions have been skillfully and adequately drawn. Had the repeated test been placed not so far apart in the recurring test periods, or had the specific tests been greater in number, the continuum of tobacco effects might have been more adequately presented. But this was evidently a laboratory difficulty of which any laboratory technician or experimenter is aware. Withal, the work is well executed and will undoubtedly take its place in the scientific literature upon the subject.

ROBERT L. BATES,
Virginia Military Institute.

MAURICE URSTEIN. *Leopold and Loeb.* Pp. 132. Chicago Medical Book Co., Chicago, 1924. Illustrated. $2.50 (paper cover).

A re-arrangement of the author's descriptions of what he terms *catatonia* (what Bleuler terms schizophrenia) is given from his previous monographic publications and incidents from the court-room behavior and developmental life of the two murderers used to show how they were presumably driven by the same psychosis.

Aside from the title Leopold and Loeb are very incidental figures. No attempt has been made to relate the case to psychoanalysis.

WILLIAM MARCUS TAYLOR, A.B., PH.G., B-P.D. *The More Abundant Life, or Bio-Psycho-Genetics.* Brandon, Nashville, 1924. Pp. lvi + 195.

I am not very certain that I have been able to find what this is all about. Mr. Taylor has not only hitched his book to the stars, but to about every philosopher and scientist of renown that the world has known. The climax seems to be reached, however, in announcing a series of twenty books on Bio-Psychology. These books are used as texts in the correspondence courses given—for a fat fee—by the Taylor School of Bio-Psychology, Inc., Chattanooga, Tennessee, upon the satisfactory culmination of which the Doctor's degree is awarded. To think that there are twenty more like this!

One gets the impression that this is the book for which the ages have longed and for which the gullible will fall. "Dr." Taylor takes no chances with the family romance and dedicates the book to his father, mother, and wife—a dubious honor.

DONALD A. LAIRD,
Colgate University.

EDGAR ARTHUR SINGER, JR. *Mind As Behavior.* R. G. Adams & Co., Columbus, Ohio. 1924. 301 pp.

Behavior to many students of life and mind means mechanism only. To this psychologist, mathematician and philosopher mechanism is true, but not the whole truth. The book is a series of fourteen essays, some previously published, and here offered in a single volume under two parts. Part I is given to an exposition and defense of the thesis that "mind is behavior." The author holds that mind is an observable object, just as objective as the datum of any science. Pursuing a purely empirical method he concludes that the "criterion of mind constitutes its definition" (p. 79); mind is what it does. Mind is a category invented to denote a quality of behavior, namely, resourcefulness. "If one being can accomplish a given purpose in $(n + 1)$ types of situation, another in but n of these, we shall call that in which the first is better equipped than the second a faculty of mind" (p. 82). ". . . . . . it is the difference in reaction to different qualities of stimulus that defines qualitative differences of sensation" p. 95). Accepting a distinction between mechanism and life, though denying that there is a chasm between them, he holds that just as a wave moving through a material medium is made up of the parts of the medium, so life is related to the body-mechanism. However, these wave-like forms are not "wind-tossed, but rather purpose-drawn." Life and mind have the same denotation, but their connotation is different. Purpose, too, is objectively viewed. It is the "average common result of a type of act" arrived at by an empirical study of statistics. It presents itself in the end as a measured probability (p. 68).

Part II undertakes the ambitious philosophic task of showing that a thoroughgoing empiricism may be idealistic.

Professor Singer's study is most stimulating. Exhibiting a clear and thorough grasp of his problem, he moves to his conclusions with an originality and subtlety which combines accuracy of mathematical and scientific knowledge with refinement of literary style. Whether he succeeds in establishing his main theses is open to question. The reviewer is still in doubt upon at least two points: (1) That life and mind may be adequately conceived on the analogy of a wave-impulse moving through a material medium; and (2) that purpose is fully explained as a measured probability statistically arrived at.

EDGAR PIERCE. *The Philosophy of Character.* Harvard University Press, Cambridge. 1924. 435 pp.

Keenly aware of the metaphysical implications of psychology and convinced that there must be some truth in both the mechanical and teleological views of nature, the author enters upon a critical analysis and synthesis of the fundamental, working concepts of physics, chemistry and biology. The aim of this study is to set up an "intelligible basis" for the preservation of the reality of purposive action and at the same time "give full value to the influence of bodily changes" on thought and action. Materialism cannot explain mind—and mind must be accepted as real to save science from meaninglessness; metaphysical dualism is rejected because of an impossible epistemology; therefore reality must be a multiplicity of monads differing in degree of mentality. (Cf. Leibniz, Royce, James Ward, *et al.*) Accepting Hume and Broad's conception of cause, viz., causation is probable expectation, the laws of science are held to be not necessary or fixed descriptions of nature, but approximations. Reality is essentially dynamic. Mechanism is true in the sense that nature has habits; nature's activity tends to become relatively fixed. "Reality is purpose becoming fixed in habit" (p. 402). The processes in the inorganic world are identical with organic processes; these processes differ in degree, not in kind.

Evolution began with an "unrest of mind;" the primeval mind-stuff wanted to do something. Through an inherent energy mind organized itself. By means of trial and error, and experimentation, in a somewhat Bergsonian fashion, our present universe evolved. A mechanical reality might be expected to develop in a straight line; not so minds. They must make progress (and progress means nothing without purpose) with the aid of experience. The body-mind problem, insoluble on any other theory, is thus dissolved. But how do the monads of the human organism fuse together into a unity? Rejecting the arch-monad theory (Leibniz) and the soul theory (McDougall), the author follows a suggestion by James and holds that they "form a whole because they are never separated" (p. 151). "Consciousness is the unity of a certain combination of elements, of monads, which we call protoplasm" (p. 407).

The reviewer confesses a good deal of sympathy with the aim, method and conclusions of this work. Clearly, of course, the scientific evidence on many points is so evenly balanced that choices are almost optional. Here, no doubt, temperament decides. Professor Pierce is correct, it seems to me, in his contention that, if we are to have a science of character, one which is to be useful in the building of a more mature and orderly society ethically, such a science must come out of a reasoned view of all the sciences—the social as well as the physical and the biological. The reader will look for a treatment of the empirical aspects of character; this the author has consciously omitted, desiring to make this problem the special task of a purposive psychology later.

EMORY S. BOGARDUS. *Fundamentals of Social Psychology.* The Century Company, New York, 479 pp.

In the *Fundamentals of Social Psychology* Professor Bogardus gives to his many readers a restatement with an enlargement of the second edition of his *Essentials of Social Psychology*. The point of view is essentially sociological. The book is divided into four parts. Beginning with a brief survey of the history of social psychology, the author devotes Part I to a psychological analysis of the individual's inherited equipment, his affective, cognitive, habitual and social nature. Part II, headed Interstimulation, is concerned with the relation of the individual to the group, and treats with keen observation such topics as isolation, stimulation, communication, suggestion, the variant forms of imitation, custom and convention diffusion, discrimination, discussion, accommodation, assimilation, and socialization. Part III deals with Groups and Interstimulation. Here the different kinds of groups, such as crowds, mobs, assemblies, publics, occupational groups, as well as group opinion, loyalty, conflict, morale, control, control agencies and products, are treated. Part IV is a study of Leadership and Interstimulation. Here originality, genius and talent, invention and discovery, the different forms of leadership, such as mental, social, prestige, and democratic, are discussed. The relation of leadership to social change and world progress is also touched upon. The author closes each chapter with a summary statement of the explanatory principles employed therein, a list of review questions, also a dozen or more challenging problems together with a selection of readings.

Social Psychology is defined as a study "of the processes of intersocial stimulation and their products in the form of social attitudes and values. It obtains its data by analyzing personal experiences" (preface). As compared with Professor Allport's *Social Psychology*, Professor Bogardus gives a more balanced view and treatment. Many readers and teachers will probably wish that the author had omitted much that is apparent and obvious in individual and social action, and made room for a more rigorous psychological treatment of the primary drives and motives in social behavior. Among other new chapters the reviewer finds those upon isolation, stimulation, accommodation, assimilation and socialization especially appropriate and well done. The book is clearly written and in attractive style. It should find a hearty welcome among students and teachers who wish a social psychology with a sociological emphasis. I know of none better.

WALTER S. GAMERTSFELDER,
Ohio University.

G. M. RUCH. *Improvement of the Written Examination.* Scott, Foreman and Co., Chicago. 1924. pp. x + 193.

Perhaps the weakest point in the "test movement" in education has been the over-emphasis of the standardized test of achievement. Poorly constructed and unreliable tests have been marketed. Improper validation may also be charged against many of the tests which now have historical value only. Since common objectives did not exist in the sense of having been experimentally determined, many tests and measures have been validated solely on the basis of judgment of the test maker alone or some other equally inadequate criterion. The publication and sale of faulty achievement tests have had a very bad influence upon the unsuspecting consumer, namely, the public school teacher.

We seem to be learning by experience, however, the proper and desirable procedures. The situation has been corrected somewhat by the appearance of a few carefully validated achievement tests in standardized form. The most important step, however, has been the attempt to teach the school administrator and the school teacher the methods of objective test construction. The volume being reviewed is a very far-reaching step in that direction.

Dr. Ruch's book brings the results of experimental studies of test construction within the purview of the busy worker in education. The presentation is such that it should serve two very splendid purposes:

1. It should serve to develop a critical attitude in the mind of the

teacher toward any sort of test or examination. It provides the criteria for evaluating examinations of the traditional essay type which criteria are equally important when applied to the brief objective test or the standardized tests which are being marketed. The book is not intended to discourage the use of tests now in published form but for those who will give the volume careful study it does provide the means for discriminate choosing.

2. The person who gives careful study to the contents of the book will find that his concept of educational measurement has become more refined. It is equally true that the reader should be able to construct fairly reliable measuring devices in the form of new type content examinations after mastering the fundamental principles laid down and following the samples given as a guide to form.

Educational measurement must ultimately depend upon a careful analysis of general educational objectives as well as analysis of specific subject matter and skills to be imparted. Such work is under way but is by no means complete so that the hope of educational measurement in the immediate future seems to depend upon the training of prospective teachers and teachers in service in the ways and means of measurement. Many progressive school principals will find that Dr. Ruch's book will serve as a splendid volume to be put in the hands of his teachers to be used as a basis for study in teachers' meetings. He will find the facts of test construction crowded into a small volume so that elaborate details will not require attention. He will find a wealth of illustrative material so that he will not be faced with any lack of concreteness.

There are points of criticism as is true of any published volume. Many teachers will have difficulty with the last two chapters in the book. The person who has had little training in statistics will find the material a bit beyond their comprehension. Fortunately the work is so presented that the reader may get the gist of the methods of objective test construction from other chapters and may simply omit these two from their reading. The reviewer feels that Dr. Ruch has unintentionally omitted a reference to Rugg's article on Statistical Methods which appeared in the Twenty-First Year Book.

It is to be regretted that attention has not been called more specifically to the fact that the pupils are the focal point in any measuring program. Any measuring device may be mechanically perfect but it is not adequate until it has taken into account a number of factors which the subjects themselves present. Some attention has been given to the need for improvement of measures through use, that is, actually trying them out on a sufficient group of subjects and revision to make the measures harmonize with the findings of experience. These latter criticisms are really not criticisms of Dr. Ruch's book.

They are criticisms of what appears to be current philosophy in the field of measurement. Failure to get the subject's point of view, failure to carefully observe the test or measure in action in the light of the reaction of the subject and failure to revise measures in the light of new findings are criticisms which apply to the vast majority of measures in use.

The author has indicated his purpose in the preface of his book by the statements that "It is hoped that the pages include enough of significant work which has been done thus far in developing objective methods to provide a bird's eye view of the aims and progress of these efforts" and "Every effort has been made to provide concrete examples and illustrations of possible examination techniques. . . . . "The reader will find that the author accomplishes his purpose in this volume and does it in a very direct manner. The beginner in measurement will find the book stimulating and helpful. The student specializing in test technique and measurement will find the volume a necessary addition to his library.

GLEN U. CLEETON,
Carnegie Institute of Technology.

THOMAS VERNER MOORE. *Dynamic Psychology: An Introduction to Modern Theory and Practice.* Lippincott, Philadelphia. 1924.

The book is divided into six parts. The first, entitled, The Analysis, of Mind, briefly sets forth the nature of psychology and the roots from which it sprang. The science is defined as "the science of human personality." The other chapters in Part One deal with Consciousness, The Unconscious, Dreams and the Unconscious, Methods of Investigating the Unconscious, and the Classification of Mental Processes. Part Two, entitled, Stimulus and Response and Human Behavior, treats reflex actions and tropisms in the usual manner. Part Three, dealing with Human Emotional Life, presents an interesting variation in the shape of a pageful of sketches showing facial muscles of "expression" in action.

Part Four, entitled, The Driving Forces of Human Nature and their Adjustment, contains chapters on Instinct and Impulse, Desire, The Conflict, Psychotaxes and Parataxes, Compensation and Sublimation. Parts Four and Five (the latter entitled, Psychoanalysis and Psychotherapy), are the sections which the author stresses most warmly, as should be expected from the title of the book. They are written in a clinical vein, being outgrowths of the author's experience as a medical officer with the American Army in France and as Director of the Clinic for Mental and Nervous Diseases at Providence Hospital, Washington.

The author's position on Instinct may be seen from these extracts: "Impulses are the real psychological elements in instincts;" an impulse being defined as "a tendency that we experience, in the presence of an actual opportunity, to make use of any one of our human abilities. . . . . There are just as many impulses as there are human abilities. Instincts are merely groups of impulses or desires to which popular parlance has given names. In danger the 'instinct of self-preservation' is called into play. This means nothing more than that every human ability that can help to extricate one from the danger is called into action. . . . . Valuable as would be the study of those groups of impulses in detail to which popular psychology has given names, we must refer this study to social psychology to which it more properly belongs."

In dealing with the mechanisms of emotional life ordinarily known as defense reactions, compensation, sublimation, and so forth, the author coins a word, psychotaxis, the root taxis referring to the tendency of the mind to adjust itself to pleasant and unpleasant situations. Psychotaxes may be positive or negative, the latter being the most common. Abnormal forms of psychotaxis may be called parataxis. Parataxis that are discussed in detail are Depression, Anxiety, Defense.

In the section devoted to Psychoanalysis and Psychotherapy, the theories of outstanding psychoanalysts are discussed, a chapter being devoted to Freud, Jung, Adler and Adolf Meyer, each. In a chapter on The Technique of Psychotherapy the author gives detailed directions for conducting examinations and treating cases. A complete scheme for a mental examination is given. Numerous cases are cited, which will interest the general reader as well as the psychiatrist.

Part Six, Volitional Control, begins with a criticism of James' position that kinaesthetic images are important if not indispensable. Numerous investigations are cited which favor the doctrine of imageless thought. The author finally concludes: "the act of will is a mental element with its own quality, intensity and duration." After a chapter on the Pathology of Voluntary Action comes a chapter on the Freedom of the Will, in which the affirmative position is supported by the familiar appeal to introspection supplemented with arguments from Ostwald. The author manifests his theological interests again in the concluding chapter which he entitles, The Soul. After marshalling evidence for the existence of the soul from biology and philosophy he concludes his dialectic: "If, therefore, there must be some substrate of conscious process, something which is active when the mind is conscious, and if this cannot be a material substance, then there must be a non-material substance, that is to say, a spiritual substance or soul. Philosophy in this way confirms the conclusion of biology."

The teacher of psychology who examines this book as a possible elementary text will miss formal treatment of Sensation, Perception, Imaging and other intellectual processes. Probably the book would be most serviceable in an advanced course in Abnormal Psychology. There are a large number of terms that would be unintelligible to the ordinary undergraduate, though the author has partially compensated for this by providing a glossary.

There are a few errors that should be corrected: On page 23 the statement is made, "according to the Freudian concept mere analysis suffices for clearing up mental difficulties," though Freud's dependence on *reeducation is truthfully explained on page 258. On page 28 is* found "mental events is set." On page 38, Jung's should read Jung. Pain is persistently put in opposition to pleasure in Chapter I, Part Three. The statement is even made: "There is no sense organ for . . . . pain."

The reader is impelled to maintain a kindly attitude toward the book by reason of the broadly tolerant spirit in which the author deals with contentious matters. For example, in the criticisms of Behaviorism there is no raillery or invective. "It is difficult to understand this denial of consciousness without an insight into Behaviorism as an outgrowth from animal psychology. With this foundation, however, and keeping in mind the natural tendency of some personalities to all embracing monistic concepts and sweeping denials and affirmations, and not forgetting either the delight of the radicals to shock the sensibilities of the conservatives, and the craving of every man to bring forward something new and startling, we may understand the 'psychology' of the Behaviorists though we may have serious misgivings as to the solidity of its logical foundations. . . . . It may be difficult to study our inner, mental life, but it is undoubtedly a field of investigation and a field of investigation which has long been termed by *the name of psychology*. This *inner mental life is of interest to* many investigators, and they have every right historically to term this science of our inner mental life psychology. It is impossible to investigate everything in our mental life by objective methods for this inner experience is far richer than its manifestations in actions or reactions that can be the objects of an external observer's experience."

"Nor has Behaviorism been able to attain its goal and predict and control human behavior. A pure Behaviorist would have little place in a psychological clinic or the schoolroom or the Juvenile Court, etc. Whenever one wishes to understand any of the real problems of mental conflict, or penetrate into the real causes of the difficulties of life, one has to obtain introspections from the patient in trouble. His reactions alone will not give the insight into his personality that is necessary in order to give him the help he needs. Psychology should enable us to solve the difficulties of the human race as well as to investigate the curve of learning in white rats, dogs, cats or human organisms."

The work is broadly conceived on a philosophical groundwork, manifesting at the same time warm human interests and scholarly regard for facts. To the professional psychologist it may furnish basis for coördinating hitherto isolated facts; to the enlightened layman it will unquestionably serve as a source-book for authoritative statements regarding psychological questions that are before the attention of the public.

H. D. Kitson,
Indiana University.

NEW BOOKS AND PAMPHLETS RECEIVED[1]

Books and pamphlets for review should be sent to James P. Porter, Department of Psychology, Ohio University, Athens, Ohio.

Annotated Bibliography Dealing with the Classification and Instruction of Pupils to Provide for Individual Differences. CHARLES W. ODELL. Vol. XXI, No. 12, University of Illinois Bulletin. University of Illinois, Urbana, Ill. Price 50 cents. 50 pp.

Beginnings in Educational Measurement. EDWARD A. LINCOLN. J. B. Lippincott Company, Philadelphia. 151 pp.

Christian Missions and Oriental Civilizations. MAURICE T. PRICE. Obtainable from G. E. Stechert & Company, New York City. 578 pp.

Constant and Variable Errors of Educational Measurements. WALTER S. MONROE. University of Illinois Bulletin, Vol. XXI, No. 10, University of Illinois, Urbana, Ill. Price 25 cents. 30 pp.

Constructive School Discipline. WALTER ROBINSON SMITH. American Book Company, Cincinnati, Ohio. 275 pp.

Elementary School Standards for the Improvement of Teaching. A. S. BARR. Edward Brothers, Publishers, Ann Arbor, Mich. 159 pp.

Growth of the Mind, The. KURT KOFFKA. Harcourt, Brace & Company, New York City. 383 pp.

Instinct. L. L. BERNARD. Henry Holt & Company, New York. 550 pp.

Introduction to Reflective Thinking. Columbia Associates in Philosophy. Houghton Mifflin Company, Boston, Mass. 351 pp.

Measuring Results in Education. MARION REX TRABUE. American Book Company, Cincinnati, Ohio. 492 pp.

Mental Growth of the Pre-School Child, The. ARNOLD GESELL. Macmillan Company, New York. 447 pp.

Mentality of Apes, The. WOLFGANG KÖHLER. Harcourt, Brace & Company, New York. 342 pp.

Mind as Behavior. EDGAR ARTHUR SINGER. R. G. Adams & Company, Columbus, Ohio. 300 pp.

Modern Theories of the Unconscious. W. L. NORTHRIDGE. E. P. Dutton, & Co., New York. $3.75. 104 pp.

---

[1] Mention here does not preclude further comment.

Nature of Intelligence, The. L. L. THURSTONE. Harcourt, Brace & Company, New York. 167 pp.

Philosophy of Character. EDGAR PIERCE. Harvard University Press, Cambridge, Mass. Price $4.00. 435 pp.

Primitive Religion. ROBERT H. LOWIE. Boni & Liveright, New York. Price $3.50. 345 pp.

Psychological Tests of Mental Abilities. A. S. WOODBURNE, University of Madras. Superintendent, Government Press, Madras, India. 229 pp.

Psychology: What It has to Teach You about Yourself and Your World. EVERETT DEAN MARTIN. People's Institute Publishing Company, 70 Fifth Avenue, New York. 248 pp.

Psychology of the Pre-School Child, The. BIRD T. BALDWIN AND LORLE I. STECHER. D. Appleton and Company, New York. Price $2.75. 305 pp.

Sigmund Freud: Collected Papers. Vol. I, 359 pp., and Vol. II, 404 pp. The Hogarth Press, 52 Tavistock Square, London, W. C. and The Institute of Psycho-Analysis.

Survey of the City Schools of Marion, Illinois. WALTER S. MONROE. University of Illinois Bulletin, Vol. XXII, No. 3. University of Illinois, Urbana, Ill. Price 50 cents. 60 pp.

Training in the Technique of Study. WALTER S. MONROE AND DORA KEEN MOHLMAN. University of Illinois Bulletin, Vol. XXII, No. 2. University of Illinois, Urbana, Ill. Price 50 cents. 66 pp.

Traité de Psychologie. Tome II. GEORGES DUMAS. Librarie Felix Alcan, 108, Boulevard Saint-Germain, 108, Paris, France. 1173 pp.

PUBLICATIONS RECEIVED FROM THE DEPARTMENT OF THE INTERIOR, WASHINGTON, D. C.

Annual Report of the Commissioner of Education, Fiscal Year Ended June 30, 1924. Price 5 cents. 32 pp.

Commercial Occupations. GLEN LEVIN SWIGGETT. Commercial Education Leaflet No. 9, November, 1924. 5 cents. 9 pp.

Coöperative Vocational Guidance. GLEN LEVIN SWIGGETT. Commercial Education Leaflet No. 8, November, 1924. Price 5 cents. 12 pp.

Coördination of Business Preparation and Placement. GLEN LEVIN SWIGGETT. Commercial Ed. Leaflet No. 10, October, 1924. Price 5 cents. 16 pp.

List of Books for a Teacher's Professional Library. EDITH A. WRIGHT. Teacher's Leaflet No. 17. Price 5 cents. 15 pp.

List of References on Vocational Education. Library Leaflet No. 25, November, 1924. Price 5 cents. 20 pp.

List of References on the Junior High School. Library Leaflet No. 27, November, 1924. Price 5 cents. 11 pp.

Recognition of Health as an Objective. HARRIET WEDGWOOD. School Health Studies No. 7. Price 5 cents. 27 pp.

School Health Supervision. HARRIET WEDGWOOD. School Health Studies No. 8. Price 5 cents. 18 pp.

Some Practical Uses of Auditoriums in the Rural Schools of Montgomery County, Ala. LILLIAN ALLEN AND CORA PEARSON. Rural School Leaflet No. 34, October, 1924. Price 5 cents. 10 pp.

Sources of Useful Information for the Teacher of Home Economics. EMELINE S. WHITCOMB. Home Economics Circular 10, June, 1924. Price 5 cents. 18 pp.

Statistics of State Universities and State Colleges. FRANK M. PHILLIPS. Bulletin No. 26, 1924. Price 5 cents. 15 pp.

Technique of Procedure in Collegiate Registration. GEORGE T. AVERY. Bulletin No. 22, 1924. Price 5 cents. 26 pp.

THE SPEECH OF FIVE HUNDRED FRESHMAN COLLEGE WOMEN[1]

SARA M. STINCHFIELD, PH.D.
Mount Holyoke College

Since the fall of 1922, standardized tests in voice and speech have been given to all entering students at Mount Holyoke College. The chief value of such tests is the ease with which it becomes possible to classify students for training. Students handicapped by speech and personality difficulties may be selected for training in small groups or individually. Students who seem to conform to a satisfactory standard may be excused from speech requirements; superior students in speech attainment may be discovered at the outset and recommended to do special work or to elect suitable speech courses.

Previous experience in various colleges seems to indicate that unless students are subjected to some examination in speech, many are graduated without having made any material progress in overcoming provincialisms, speech handicaps, personality difficulties, and even vulgarities of expression. Speech departments frequently never hear of such students. Many serious cases of lisping, stuttering, and indistinct utterance are present in every college, and personality difficulties which interfere seriously with college social adjustments are often ignored. *As yet, the majority of our colleges are doing* little to remove these handicaps.

There should be no cause for disagreement along different lines of speech work. There seems to have been a certain amount of misunderstanding, however, in the past because of the failure to recognize the dividing line between the artistic and scientific approach to speech problems. We need to know how to harmonize the two.

---

[1] Paper given at the joint conference of the New England and Eastern Teachers of Speech, at Harvard, April, 1924.

In music much the same condition exists—there is the difference between teaching a child to play the piano and teaching him to tune the instrument. Most people realize when something is wrong with the piano, and they send for the expert, the piano-tuner. I have heard of a musician who found an old square piano quite out of tune, and as the owner was willing to risk his tampering with it, he undertook to tune it. Having a good ear he made it sound very well in the key of C in which he did the tuning, but not understanding the principle of the tempered scale, he left the piano out of tune in every key involving the use of sharps and flats.

The same thing often occurs in the attempt to correct minor speech defects. A person trained along interpretive lines, like the pianist, may modify poise, appearance and speech reactions of a speaker so that under controlled conditions or in some situation the speech is correct and even excellent, but various faulty temperamental difficulties in the individual have not been even suspected. Many types of speech defects with "introversion" or emotional imbalance need the scientific approach for their treatment, as does the successful tuning of the piano, otherwise one may obtain an accentuation, rather than a correction of the difficulty. For instance, in stuttering a person is often referred to various unscientific schools which "guarantee a cure." Of course no reputable physician ever *guarantees* a cure. I have just received a letter from a lad protesting strongly against such unethical methods. He writes: "Is there any standard method for the correction of stammering, which is at least partly reliable? There are so many relapses, after one has taken various *guaranteed cures*, that one loses hope!"

The modern scientific methods of correcting speech disorders have followed the principles of physical medical treatment, i.e., not to treat *symptoms*, but *causes*. Take rheumatism for instance—our grandparents rubbed oil on lame joints to secure relief from pain, but there was no known *cure* for rheumatism. Nowadays we find substances within the joints which produce rheumatism and the causes are often found in autointoxication produced by such conditions as diseased tonsils, decayed teeth, etc. Physicians do not treat *headaches*, though there are plenty of patent medicines which claim to do so. The scientific medical man seeks the cause of the headache and attempts to remove the symptoms by changing the causal conditions.

The purpose of our speech examination is selective; the plan being that students handicapped by ineffective voice and speech or personality difficulties and maladjustments may receive the necessary training; that students who conform to a satisfactory standard may be excused from training; and that special talent may be encouraged.

To secure a speech index or speech rating, the standardized Blanton-Stinchfield Speech Measurements (6) were used, which include an estimate of personality, of articulation of all consonant and vowel sounds in the language, of speech and voice, in the form of a series of objective and subjective measurements. The report of our findings is largely in tabular form with a brief explanation of each table.

Table 1 gives the distribution by groups of the entering classes of 1922 and 1923, with percentage of students in each group.

The striking fact in this study is, that although about one-third of the students last year and nearly half of the students this year passed the speech requirement, at least one-sixth of the Freshman class need definite corrective work, individually or in small groups. This makes a group of more than fifty in this year's class.

Besides these' another considerable portion of the class, from one-fifth to one-fourth need speech training which is postponed to the Sophomore year. This group consists of less urgent cases than those in the corrective group. There seems to be no doubt that they should receive training before being allowed to graduate.

It is a source of great satisfaction however to find an equal number (one-fifth to one-fourth), who seem to have sufficient native ability to justify us in encouraging them to develop

TABLE 1

GROUP		YEAR 1922		YEAR 1923	
		Number of students	Per cent of class	Number of students	Per cent of class
			per cent		*per cent*
I	Required Freshman work in speech correction	39	18	53	16
II	Required Sophomore speech work (ineffective speech)	42	19	80	24
III	Excused from speech requirement (average group)	75	35	144	42
IV	Recommended to elect speech because of good standing in the tests	60	28	63	18
Totals		216	100	340	100

Total number of students for both years, 556.

TABLE 2

Frequency table showing types of speech difficulties for which students are placed in Groups I and II (required Freshman and Sophomore speech classes), totalling 133 students. 1923–24

	PER CENT
Oral inactivity and ineffective speech	41.0
Lisping (on s and z sounds)	17.0
Faults of voice, rhythm and breath control	25.0
Stuttering, hesitation or other indications of nervousness	6.0
Foreign accent	2.3
Letter substitution	0.7
Mixed (more than one type of speech difficulty)	8.0
Total	100.0

their talents by elective work, as voice, personality and social adjustment in these cases seem to be of the normal, well-balanced type.

The total number of entering students, September, 1923, was 340. Of these 16 per cent were classified in required Freshman corrective group, and 24 per cent were classified in required *Sophomore speech classes* (see table 1).

The largest group in this table consists of those who show general oral inactivity and ineffective speech, defects which seem to be based upon the general constitutional make-up of the individual. These cases have been particularly difficult for teachers in speech classes, being reported as "uninteresting" and "unreceptive" or "unresponsive" to training.

> Speech disturbances disappear when the individual feels free to express himself; that is, when he can give and take emotions in a normal way. One must remember that the neurotic symptoms are, as it were, a morbid gain for the patient, and unconsciously he refuses to give them up. Chronic stutterers, like other chronic neurotics, make use of their stuttering to escape from all kinds of difficulties. Stutterers, the world over, stutter most when they are confronted with disagreeable tasks, and a great many of them who have remained "cured" for years, relapse when confronted with a painful situation. So that our treatment consists in showing the patient how to face reality. There is no royal road to the treatment of speech disturbances. It is the most difficult neurosis that one has to manage. There are very few *cures* effected among the adult chronic stutterers (1).

> All motor tics and habit spasms have an unfortunate effect on speech, as the motor activities tend to become less rhythmical and induce a condition of nervous exhaustion. Certain tics because they interfere directly with the free use of the vocal and speech mechanisms, must be specially considered by the speech worker (2).

> Many of these tics have their foundation in some slight physical disability, which should be removed. Some, unfortunately, have their foundation in psychic disability, the removal of which is a longer and more difficult process. All, whether of physical or of psychic origin, have a definite psychic value. *The organism does not continue useless movements.* They may be apparently useless, may even seem so to the person who has acquired them, but that they are emblems of some need or serve some purpose in the emotional life of the person cannot be doubted (3).

Although the experiment is relatively new at Mount Holyoke, we feel justified in saying that the more severe types of speech defects, emotional imbalance and lisping, sometimes give the better prognosis, because in such cases the student more readily recognizes her handicap, and exerts herself to overcome it. Hence, what sometimes seems an utterly hopeless case, may turn out to be most hopeful, while comparatively mild cases, may make little if any improvement. This applies particularly to hysterical stutterers, in whom surprisingly good results are often obtained in a short time, while seemingly mild stutterers show little improvement after the expenditure of much time and effort on the part of the teacher and the pupil. This does not hold for chronic stutterers, or for many cases of speech disorder where there are severe organic and functional factors in the background, such as oral and nasal malformations, epilepsy, infantile paralysis, etc.

It will doubtless be surprising to many persons to learn that among the highly selected group of girls who are admitted to college, we find such a variety of types of speech difficulties, all of which are recognized as more or less serious handicaps in post-college days. Considering the great stress that is placed upon appearance, address and behavior by employers of women, it seems a very great misfortune for our girls to fail of success because of a remedial speech difficulty.

A striking illustration of the need of speech training during college days came to the attention of the writer recently, when a high school teacher of English was referred to us. Her speech was of the oral inactivity type in which the tongue contacts for various consonants were inaccurately made, and there was a disagreeable lateral lisp present. No one had ever called the attention of the girl or her parents to the presence of this speech defect. It was only when she began to teach that she realized that she possessed a speech handicap.

Another instance came to our attention recently in consultation with a dentist who has suffered a lifetime from feelings of inferiority due to congenital cleft palate. This condition was allowed to go uncorrected for several years with the result that the youth became embittered, sensitive, brooding and morose to the extent that in adult manhood he cannot throw off these moody tendencies, even though the condition which originally produced them, has been removed. He recently refused to appear before two distinguished dental organizations, where he had been invited to demonstrate his newly invented dental appliance—his refusal being based on fear of failure or ridicule, even though he now has a pleasing personality and a rich resonant voice quite free from nasality.

> If the speech defects are in reality symptoms of an unwholesome temperamental condition, attempts to cure the symptoms without reference to the temperamental factors are found to be ineffective. It is obvious that recurrence of the symptoms is to be expected. Occasionally the speech trainer may be fortunate enough to *resolve the conflict*, and so bring about permanent relief. Temporary alleviation which relapses as soon as the patient dispenses with treatment indicates how great is the force of suggestion, *how largely the speech defective depends upon others for the stamina necessary for proper speech* (4).

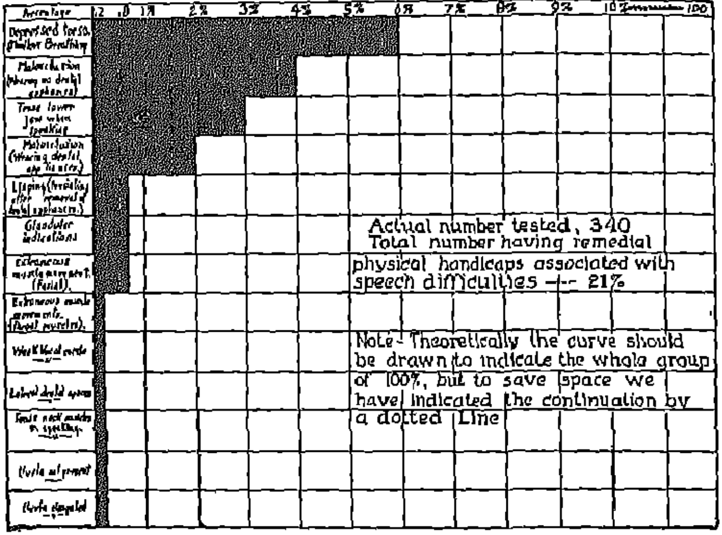
Moreover adequate provision seems to have been made for corrective work for most of the physical defects which do appear. Access to the health and medical reports of these students was possible through the coöperation of the department of physical education and it seems clear that very few students with serious physical defects are admitted.

There is a striking difference between the findings given in graph I and graph II, in which mental and physical difficulties are indicated.

The obvious conclusion resulting from a comparison of these tables is that while 21 per cent show physical defects, nearly twice as many students are below normal in the ability to adjust themselves to their social environment. (By normal, as here used, we mean those people who are classified as belonging to the active-controlled, socially adjustive type.) Twenty-one per cent is a much lower proportion than one would expect to find in a group of this size. The college group is obviously a selected group (graph I).

Graph II (behavior types and psychic difficulties) indicates that 36 per cent need attention.

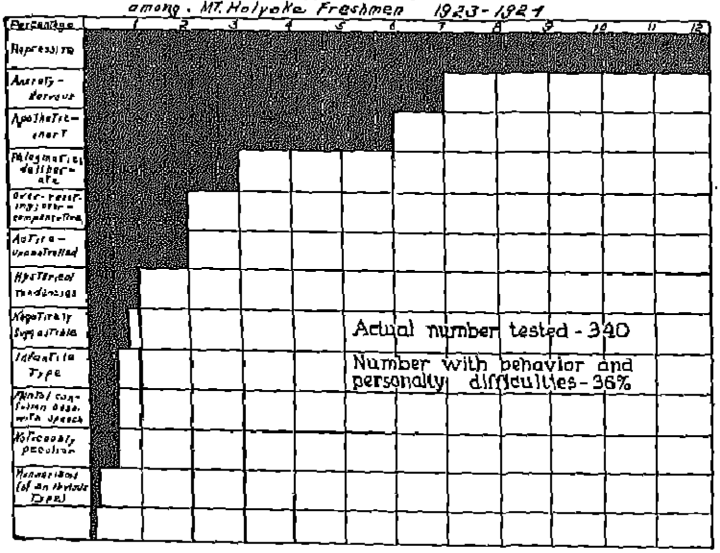
Graph I —- Showing Physical Conditions Associated With Difficulties Among Mt. Holyoke Freshmen —— 1923-24



Actual number tested, 340
Total number having remedial physical handicaps associated with speech difficulties —- 21%

Note - Theoretically the curve should be drawn to indicate the whole group of 100%, but to save space we have indicated the continuation by a dotted Line

Graph II Showing Behavior Types and Psychic Difficulties among . Mt. Holyoke Freshmen 1923-1924



Actual number tested - 340

Number with behavior and personality difficulties - 36%

> Adjustment to the exigencies of life is the great and fundamental mental law on which all human intercourse is based, and all human happiness. The child cannot dominate the world. If he could it would not bring him happiness, because it would not bring him friends. The earlier he is taught the simple practice of looking at his needs from the point of view of society, desiring for himself only those things that would be reasonable if demanded by the whole race, and accepting those things that it were not well should be changed, the greater will be his opportunity for normal emotional life, and healthful spiritual development. Normal speech is only present where these two attributes are present (3).

In the Mental Hygiene Primer, Dr. Pratt writes:

> Although men and women in years, many persons because of their mental make-up, for which early parental influence and environment are chiefly to blame, habitually carry over into adult life many of the dependent habits and customs of thinking of their childhood, only to find that reality demands a life adjustment based on mature mechanisms.
>
> The whole mental hygiene movement is rapidly getting back to its starting point—early childhood—and is dedicating a large portion of its program to this period. During this flexible and formative age sound habits of mental health can best be instilled and faulty ones corrected.
>
> Mental hygienists are stressing one great point, namely, that in most cases of nervousness, in many cases of delinquency, in some cases of insanity and in almost all cases of child behavior or conduct disorder, *the trail leads inevitably and directly back to the home and the parents.* And this fact operates in just the same fashion and with almost as much vigor and frequency among families of the well-to-do as it does in the tenements (5).

Social and economic maladjustments, behavior problems, psychopathic reactions ranging from feelings of inferiority to suicidal tendencies, often have their roots in mental difficulties of individuals who may influence large groups. A peculiar mental "twist," warped judgment, or feelings of inferiority due to repeated failure, may be the explanation of incipient and mild behavior disorders among college girls, particularly those of a repressive, introvert type. While there is much to be said in favor of a limited amount of "New England reserve," there are undoubtedly cases of anxiety neurosis which result from severe repression of the normal play life, or from severe criticism by the elders of comparatively mild offenses in adolescence. Graph II suggests rather a large per cent of such cases among our entering college girls.

We sometimes find an "apathetic-inert" girl, who is passive rather than active in her attitude toward life; or she may be of a sophisticated, "poseur" type Her reactions seem characteristic of a physically passive, rather than of an active individual—not unlike the frail, exotic Burne-Jones ideal of half a century ago This is not the type of girl who wins prizes in the athletic contests of today. She may be a good student, and yet suffer from innumerable minor "ailments."

The "phlegmatic or deliberative" type differs from the above in mental and in physical characteristics. This type of student inclines to be faithful and plodding, but rarely takes high honors. Instead of being erratic, she is steady, industrious and fairly regular in habits, but less artistic, more easy-going and less easily disturbed than the former type. She may not fail in her work but she is inclined to be slow-moving, lethargic, deliberate or procrastinating.

The "over-reacting or over-compensative" type often results from the desire to conceal some real or fancied inferiority. The girl is anxious to be well-liked, and being unable to attract friends easily enough in ordinary ways, adopts bizarre, fantastic or clownish tricks and mannerisms which make an appeal to her group, because of their unusual character, novelty and seeming originality.

There are always a few girls who find their way to college, who have been singularly isolated from playmates of their own age, or who have developed a feeling of social inferiority as a result of some real or imagined family stigma. The impression made upon the memory of the growing girl seems well-nigh irradicable. Thus, the daughter of a divorcee, in one instance, appeared to be a sophisticated "flapper" of the cosmopolitan type, world-weary at eighteen, "bored" by most of the youths of her acquaintance. In reality she was a girl of unusual mental endowments, who stood high in college work, being witty and attractive socially, despite her affectation of "boredom." She was distrustful of men and wary of their attentions, because of her knowledge of her mother's unhappy marriage. When some of her unethical reactions were discussed with her by an older woman whom she liked, and her responsibility for influencing the standards of her group was brought to her attention, she became coöperative, responsive and a valuable asset to the group. Though her energies seemed sometimes foolishly directed, there was a definite change in attitude, and in the influence which she exerted. The active-uncontrolled girl is often one of the most useful, socially, and important to the group in which she belongs. Much can be done to control, curb or direct her tendencies, under right conditions.

The "hysterical" girl is more commonly discovered by the room-mate of the offender than by members of the faculty. She constitutes a serious problem, as she readily gains the sympathy of the uninitiated by her temper tantrums, crying spells, "nervousness" and other neurotic symptoms. She must have an audience for her "attacks," so she frequently chooses the time when she actually disturbs the work of a large group of students, who attempt to minister to her needs by giving the desired sympathy. In one group it was quite common for a girl accused of indiscretions and irregular conduct, to retaliate by hysterical crying "fits," in which her violent efforts to secure sympathy served to distract her accusers. In this way she sought to avoid the deserved rebukes of her companions who were entirely justified in demanding that she conform to higher standards of conduct. The reactions of the hysteric can best be dealt with by an experienced person who knows when sympathy should give way to firmness, decision and judgment in dealing with the matter.

The "negatively-suggestible" person is commonly known as "stubborn." Such a girl reacts to suggestion by going directly contrary to the proffered advice, or she makes a negative adjustment in which her parents lose her confidences. Teasing on the part of older brothers and sisters, relatives and even by parents may develop such a latent tendency in any child. It is not that one "takes after" one or the other parent in this matter! It is much more likely to be an acquired characteristic. This stubbornness is a difficult thing to overcome in adult life, though the knowledge of the danger may assist one in guarding against it.

Tendencies to react in any of the above ways are present in many of us. The reaction may depend upon education, training, environment or immediate stimulus, decidedly more than upon inherent characteristics, social inheritance or ancestry. Certainly no young person should be graduated from our colleges without having these tendencies, when they exist, brought into consciousness, in order that the individual may have help in solving her difficulties, and in overcoming handicaps before they become a part of her habit equipment, personality and fixed attitude towards life.

SUMMARY OF RESULTS

1. From the standpoint of articulation the most frequent causes of failure are the provincialisms and local dialects in New England and the middle-Atlantic groups, the inverted "r" of the middle West and the quality of the Southern vowels.
2. Many minor defects were discovered which indicated personality difficulties, emotional imbalance and ineffective social adjustments (graph II).
3. A high percentage of inaccuracies on s and z sounds and their combinations were found (table 2).
4. On the whole, students from private schools did better than those from public schools and those who had had a little speech training seemed to pass the tests more easily than those who had received none.
5. Negligent lisping or oral inactivity seems more common than organic lisping (this latter being due to some physical defect in the mouth or in the dental condition).
6. Of the nine Freshman students sent home at mid-semesters, for failure in college work, four, or nearly fifty per cent, were classified at entrance to college in the speech correction group.

7. The speech tests give us an opportunity for observing emotional disturbances, feelings of inferiority, peculiarities of personality, nervous instability and so on, which if left uncorrected may interfere seriously with success in college and afterwards. Our chart (graph II) shows how much more common these mental disturbances are than are physical defects in the college girl. The latter have been provided for, for many years by the corrective work of the physical education department. We hope to do an equally effective job with the mental peculiarities.

REFERENCES

(1) Brill: Quar. Jour. Speech Educ., April, 1923, pp. 129–135.

(2) Brill: *Fundamental Conceptions of Psychoanalysis*. *Harcourt*, Brace Co., New York, 1923.

(3) Blanton, S., and M.: Speech Training for Children. Century Co., New York, 1920, p. 77; p. 144.

(4) Downey, J. E.: The Will-Temperament and Its Testing. World Book Co., Yonkers, N. Y., 1922, Chap. XV.

(5) Pratt: Mental Hygiene Primer. Publica., 42, Mass. Soc. for Mental Hygiene, Boston, Mass., 1923.

(6) Blanton-Stinchfield Speech Measurements. Pub. by C. H. Stoelting Co., 3037 Carroll Ave., Chicago, Ill.

Additional References

Jones, Daniel: An Outline of English Phonetics. G. E. Stechert Co., New York, 1922.

James, Henry: The Question of Our Speech. Houghton-Mifflin Co., Boston.

Klinghardt: Ubungen im englishen Tonfall. G. E. Stechert, New York, 1923.

Birmingham-Krapp: First Lessons in Speech Improvement. Chas. Scribner's Sons, New York, 1922.

Doll, E. A., and Leatherman Z.: Maladjustment Among College *Students*. *Jour. of Appl. Psychol.*, Dec., 1924.

Burnham, W. H.: The Normal Mind. D. Appleton Co., New York, 1924.

WHEN IS A MAN INTOXICATED?

H. L. HOLLINGWORTH

Columbia University

"When is a man baptized?" was once a provocative question. When he has been immersed, or only when he feels a holier impulse, or not until his fellows can testify to his good works? In somewhat the same captious vein the question has more recently been, "When is a man intoxicated?" When he "feels a misery in his breast?" Or only when the neighbors complain of him? Or not until a technical expert discovers a deviation in his reaction time or alcohol in his blood stream? The practical question, as to what does or should constitute the legal criterion, does not concern us here. It is obvious that there are degrees of intoxication, and that the psychologist's notion of it may not accord with that of the policeman.

The question we raise, instead, concerns the way in which various usable criteria are related to the different degrees of toxic influence. Excluding the criteria of the chemist and the physiologist, at least three types of indication are experimentally available in a psychological study of the nature of any drug effect. These may be briefly described as follows:

a. Subjective or introspective report. The individual's verbal communication of the state of his feelings, his array of sensations, impulses, inclinations, and other experiences.

b. General conduct or overt behavior. The individual's attitude, appearance, gait, gesture, facial expression, motor control and the like, capable of report by another observer or witness.

c. Technical measurements. Accurate determination, under controlled conditions, of changes in capacity or variability in the performance of standardised tasks, functions or processes.

In all three ways—introspective account, overt behavior, and technical measurement—may be sought symptoms of the effect of a drug, or, indeed, of any other influence. Of these three methods, which is the most accurate, the most sensitive? Which type of symptom detects the smallest change or betrays the least influence? For one thing, the answer to this question would throw considerable light on conflicting expert testimony in connection with drug effects. Nor is it likely that an *a priori* reply can be anything but a guess. For this reason it seems worth while to report an experimental inquiry into the facts.

Under conditions more fully described in my recent papers on "The Influence of Alcohol," six male subjects were studied for a period of two weeks. During this time they were given, according to a systematic schedule, various dosages. On some days no dose of any kind was given (blank or normal days). On others varying amounts of a control dose were administered (control days). On other days amounts of alcohol ranging from 39 to 78 cc. were given at mid-day, in a brew equal in quantity to the amounts of the control brew.

Each subject kept a "personal data book," in which each day he recorded his general state of well-being, any unusual feelings or sensations, his evening activity, and the amount and quality of his sleep that evening. Five judges (with psychological training) were given these reports, shuffled in random order, each man's report for a given day being on a separate slip of paper. The judges were told that on some days the man had alcohol, on some days the control and on some days nothing. They were asked to sort the slips into these three piles, according to their best judgment of the significance of the reported introspections.

On the other hand, three observers who were present in the laboratory with the men throughout the whole time of their work had every opportunity to note their conduct, conversation, appearance, their attitude toward one another and toward the experimenters, and observed them during the test procedures, but without making comparison of records. At the end of each day, each observer (two women and one man) reported any unusual signs or symptoms noted in any of the subjects, or else wrote them down as normal or usual in every way. Five judges (those previously described) now arranged slips bearing the separate reports (by all observers for each man for each day) into three piles, according to their judgment as to whether the observer's report indicated alcohol, control, or normal day.

At the same time the objective measurements in an array of tests were secured, and these have been presented in other papers. For the present we are merely interested in the general results of these tests and the question of the degree of drug influence which they were able to detect. In this respect we are now able to compare the three modes of report just described.

By way of illustration of the type of material borne by the slips the following samples are shown, chosen at random. Since there were six subjects and five judges, there were for each experimental day 30 verdicts based on introspection and 30 based on reported conduct. We have distributed the verdicts for each subject, and then combined the results for all six subjects, thus giving the general effect of the day's judgments, based on the introspection or the conduct reports for the day in question. These resemble somewhat the task of a judge or jury, confronted with the defendant's statements on the one hand and the testimony of witnesses on the other hand, the testimony being limited to "observed facts," and excluding inference.

SAMPLE DATA SLIPS—INTROSPECTIVE REPORTS

11-II. Fair this forenoon, sick and nauseated in afternoon. Later in evening very disturbed, dizzy and exhausted. Imagination terrifically active. Wrote a song before going to bed. Slept heavily, usual time.

12-I. Felt lively in forenoon, cheerful and comfortable. About the same in the afternoon, but a little tired. Felt fine in the evening, and took a walk. Sleep good, usual amount.

17-III. Felt good in the morning. Some dizzy in the afternoon, which began to wear away at about 3 p.m. As usual in the evening, perhaps a little dull. Sleep good, usual amount.

10-IV. Very slight cold, few sneezes during the day. Otherwise as usual. Sleep as usual, good.

17-V. Very active in the morning, with keen desire to make better records. In afternoon slightly exhilarated, and sleepy from 2 to 4 p.m. Feeling was slightly benumbed. Lazy all afternoon. Power of nervous tension seems taken away, without which I cannot make a good tapping record. In the evening, as usual. Sleep, as usual.

SAMPLE DATA SLIPS—CONDUCT REPORTS

11-II. A. Exuberant at 1 p.m., tended to break away from routine (thus brought chair from next room, which he had never done before). Then became sick and mournful, slow, could scarcely steady self for pulse records.
B. Blood shot eyes, can't seem to see right, very slow, holds his head. Better at four o'clock.
C. Frightfully sick, so he couldn't do any tests with a quarter of his usual vim. Complains he never felt more beastly in his life.

11-V. A. Seemed extra cheerful, but no other symptoms.
B. Talked more than usual, inclined to laugh. Unsteady and said he "felt happy."
C. All right, but "silly."

14-III. A. Nothing special noted.
B. As usual.
C. Nothing special noted.

16-VI. A. Said he felt exhilarated but did not show it by talkativeness.
B. Same as usual, perhaps a little less talkative.
C. Acted as usual. No comments.

Table 1 gives the percentile distribution of the judgments for each of the experimental days here studied, when all six subjects are grouped together, and the judgments of the three observers all included. In the first vertical column is given the dosage for a given day. In the other columns are given the distribution of the verdicts, by the two methods, under the three headings.

On the whole the 39 and 52 cc. doses give results closely alike, and the 65 and 78 cc. records are also very similar. But the results for the former (called here small doses) are very different from those for the latter (called large doses). Comparison of the results can most reliably be based on the four arrays—normal days, control days, small alcohol days and large alcohol days. The following conclusions are suggested by such a comparison of the distributions of the judgments—there being in all 60 judgments for each type of day.

TABLE 1

DOSAGE	BASED ON INTROSPECTIVE REPORTS			BASED ON BEHAVIOR REPORTS		
	Alcohol	Control	Normal	Alcohol	Control	Normal
	per cent	*per cent*	*per cent*	*per cent*	*per cent*	*per cent*
Normal record, two days......	13	3	84	7	0	93
Control record, two days......	10	57	33	5	10	85
Alcohol doses, small:						
39 cc....................	60	30	10	23	10	67
52 cc....................	50	23	27	33	47	20
Average....................	55	27	18	28	28	44
Alcohol doses, large:						
65 cc....................	83	7	10	77	13	10
78 cc....................	70	10	20	73	17	10
Average....................	77	8	15	75	15	10

Blank Days are correctly judged as "normal" with less than 10 per cent error, the errors being very slightly biased toward the alcohol verdict. Judgments are about equally reliable, whether based on the introspective accounts or on the conduct reports.

In the case of the control days, judgments based on the conduct reports are not distinguishable from the verdicts rendered on blank days. But in the case of judgments based on introspective accounts, the control days are diagnosed correctly in considerably more than half the number of cases. Based on these accounts the control days are just about as correctly identified as are the "small alcohol" days, on the same basis.

With "small alcohol" doses, about half the judgments based on the introspective accounts will be for and half against the verdict "alcohol." Of the half against alcohol, about equal numbers favor control and normal. That is to say, three-fourths of the judgments agree that "something is wrong," but only half of the total judgments correctly identify the condition. Based on the conduct records, however, the verdict is against alcohol, in a ratio of about 3 to 1. The non-alcohol votes here favor normal, but the votes for normal in this case are only half as frequent as on actual normal days. Judgments based on subjective account are thus more reliable than those based on behavior record, for the small doses. That is to say, alcohol affects the feelings before it does the grossly observable conduct.

With "large alcohol" doses, the votes are overwhelmingly for alcohol, whether based on subjective account or on behavior report. And the distribution of the judgments is quite alike in the two cases.

Generalizing, we may say that, if a majority vote be allowed to decide for or against intoxication, the small alcohol doses do not produce it when the judgment is based on the behavior reports. When the verdict is based on the introspective account, the alcohol effect is just barely said to be present.

Much the same thing is to be said for the control days, except that here the behavior reports are of no value at all and control days are confused with normal rather than with alcohol days. Large alcohol days are correctly judged by a three-fourths vote, whether based on introspective account or behavior report.

During the course of the experiment, measures were made of ability in steadiness, tapping, coördination, color naming, adding, opposites, substitution, paired associates, and pulse rate. Every one of these measured traits or processes reveals the effect of the "small alcohol" doses, positively and unambiguously. This effect the behavior reports failed to reveal, and the introspective accounts disclosed it only equivocally and by a narrow margin.

In fact, in a special experiment, covering one day, in which the effect of added dosage was studied, beginning with 13 cc., four of the technical tests, viz., adding, substitution, opposites and paired associates, began to show definite alcohol effects after only 26 cc. These are the most clearly verbal, abstract, symbolic or mental of the processes examined.

It seems clear then that effects may be present which are measureable by properly chosen technique long before they are manifested in the individual's gross conduct in a normal, social, working situation, and even a considerable time before they are indicated by the subjectively reported symptoms.

The order, in decreasing accuracy, is as follows:

A. Testimony of technical measurements
B. Evidence of subjective or introspective accounts
C. Judgments based on general conduct or overt behavior

In the present experiment the third of these identifies only the large dose effects, the second correctly identifies the small dose effects, whereas the first correctly identifies effects from doses smaller than any of these here considered. In the case of alcohol, at least, the effects will need to be of that order of magnitude produced by what are here called "large doses" (65 to 78 cc.), consumed within a period of not more than one hour, before all three sorts of testimony agree.

It is at least clear, at any rate so far as alcohol is concerned, that the technique of experimental psychology is capable of detecting drug effects considerably smaller than those revealed by introspective testimony or by general observations of appearance and behavior. Whether or not "toxic effects" are present can be determined only after the method of their detection is specified. Until such specification is agreed upon or conventionalised, we may expect expert witnesses to disagree and judicial decisions to be uncertain.

Finally, it is quite possible that the value of these three modes of detection will vary with the drug concerned. That drug effects, of any kind, will be shown in gross conduct more readily than in the technical measurements, seems quite unlikely. Whether or not some drug effects show themselves more reliably in introspective account than in either of the two other methods, is at least problematic. It is also entirely possible that the method best suited for one individual is not the most sensitive for another—that autogenic reinforcement, for example, might enable one individual to maintain his technical performance and also to control his appearance and general conduct, although a frank introspective account would reveal the presence of a disturbing influence. In the present instance we have ignored such individual differences in favor of an indication of the general tendency in the case of the three modes of detection.

Definitely supporting these conclusions concerning the relative accuracy of the three methods, are the results of some experiments which I reported about twelve years ago. In this case a study was made of the accuracy of performer and of witness in judging whether a test performance just completed was "usual," "worse than usual" or "better than usual." Since actual measurements were also made, but kept from the knowledge of the judges, we have in this instance also the three techniques—performer's introspection, observation of witness, and technical measurement.

Considering all those cases in which the judgments were rendered with anything but the lowest degree of confidence (A and B judgments) the per cent correctness of the verdicts by the three methods were as follows:

A. Testimony of technical measurements.	100 per cent correct
B. Introspection of performer............	84 per cent correct
C. Judgment of witness..................	80 per cent correct

The order of relative accuracy is quite the same as in the case of the present drug experiments.

It is not my purpose to impugn the practical value of the usual testimony. It may well be that for legislative purposes the most useful criterion is to be found in the rough observations of a witness, unsupported either by introspective accounts or by technical measurements. But this method we have shown to be the least accurate and the least sensitive of the three techniques here studied. When it is concluded that "there was not the least sign of intoxication;" that the drink "did not in the slightest degree impair the mental faculties or bodily functions;" and that "none manifested the slightest symptom," it is at least scientifically, and perhaps practically important to realize that the presence or absence of a symptom is obviously a function of the technique employed in searching for it. In the present connection we have been interested merely in the determination of the relative accuracy of three such techniques.

THE CLINICAL VIEWPOINT IN VOCATIONAL SELECTION[1]

MORRIS S. VITELES

University of Pennsylvania

This paper grows, in part, out of a conference with the works manager of a metal working plant. This gentleman was faced with a production problem in three departments: (1) a press room, (2) a milling and drilling room, and (3) an assembly department. He had developed an almost naïve faith in the potency of psychological tests for solving this problem by selecting workers best qualified for each of these three types of work. In discussing the proposed tests the works manager insisted upon placing the following limitations upon the work of the psychologist.

1. That the tests be limited to between thirty and forty-five minutes.
2. That only machines and material found in the plant be used in the manufacture of the apparatus involved.
3. That the tests, when devised, be capable of administration and interpretation by a clerk in the employment office.

It is the third limitation which is to be discussed in this paper. The first impulse is to criticize the works manager for insisting upon this limitation. But it requires only a hasty glance over the literature of vocational selection to discover that he was merely insisting upon a condition which the vocational psychologists have themselves established. Again and again, both in technical articles and in advertising announcements, have these psychologists repeated that these tests can be given, graded, and evaluated by a minor clerk in the employment office. Only rarely has it been suggested that the objective scores of vocational tests are at best uncertain diagnostic criteria; that these must be supplemented by an observation and analysis of performance by a trained examiner.

---

[1] A paper read at the meeting of the Industrial Section, American Psychological Association, Washintgon, December 30, 1924.

For example, in a lengthy exposition on "Measurement in Vocational Selection: an Outline of Research Procedure," Freyd states that[2]

> it requires more training to conduct an experiment in vocational tests than to administer the tests. An economy of effort will result if the experimenter assigns the task of administering the tests to another. He should insist on the selection of a tester who is fairly intelligent, polite and tactful, and favorably inclined toward the theory of tests and scientific method in general.

In a later section of the same article he says that "the maintenance of test procedures and other measurements in vocational selection clearly demand an oversight of a carefully statistically trained and judicious-minded person." It is important to note that in neither the case of the supervisor of test procedures nor of the administrator of tests is training in psychology mentioned or even implied. Apparently, politeness and statistics are more important for successful work in vocational selection than psychology. Moreover, it is the duty of the overseer to supervise the maintenance of test procedure. In no place is it suggested (and Freyd's viewpoint in this connection is a representative viewpoint) that it may also be the function of psychologists working in industry to evaluate and interpret not only the score, but the actual test performance of the individual applicant for employment.

The insistence upon this point of view is based upon a number of assumptions.

1. The first is a statistical assumption. It says that[3]

---

[2] Max Freyd, Measurement in Vocational Selection, Journal of Personnel Research, vol. II, nos. 8 and 9, December, 1923—January, 1924, p. 377.

[3] Max Freyd, Measurement in Vocational Selection, Journal of Personnel Research, vol. II, nos. 8 and 9, December, 1923—January, 1924, p. 272.

> the test score in the practical situation is always the known variable and from it the probable degree of vocational success may be predicted on the basis of the known relationship of test scores and degrees of vocational success. According to this assumption the evaluation of the tests consists in the comparison of two series of measurements made of the subjects of the experiment, the test scores and the criterion of success.

There is no time here to criticize in detail this assumption. Those of us who are interested in the examination of children for educational purposes, for correctional purposes, or for purposes of guidance have found it impossible to diagnose and treat on the basis of objective score alone. It has been found necessary to supplement the objective data of mental measurement with the observation and analysis of performance. Mental ages, I.Q.'s, and time scores have been found to be useful in guiding the examiner, but not in directly indicating the intellectual level or specific abilities of children.

There is a qualitative aspect of mental analysis which goes hand in hand with quantitative analysis. Maxfield points out that[4] "in the use of standardized mental tests subjective judgments of the experienced examiner in regard to the qualitative aspect of the subject's reactions, whether verbal responses or other types of behavior, are significant."

The clinical psychologist has been forced by experience to place himself in the position of a physician in the matter of diagnosis and recommendation. With reference to diagnosis the psychological test may, in a sense, be compared to the various objective tests, some of them laboratory analyses, made by the physician's assistant. The physician permits his assistant to test the temperature of the patient, to read his pulse; he seeks the laboratory technician's report on the blood count, the incidence of sugar in the blood, the rate of oxidation, etc., but in practically no instance does he throw upon the assistant technician the responsibility of a diagnosis. Diagnosis involves an interpretation of the objective data—interpretation involving scientific knowledge which the assistant does not possess, and which, therefore, makes him incompetent to render a diagnosis. There are very few exceptions; the technician's report of a four plus Wassermann establishes without doubt the presence of syphilis, and in a sense dictates the diagnosis and the treatment, but even in medicine these specifics are rare, and in the field of psychology the test which is a specific diagnostic measure is still rarer, if it exists at all.

---

[4] F. N. Maxfield, The Use and Abuse of Standard Intelligence Tests in Individual Examinations. Reprinted from the Proceedings of the 48th Annual Session of the American Association for the Study of the Feebleminded, 1924, p. 21.

"Many psychologists," as Maxfield[6] wisely observes, "have had the tendency to confuse mental measurement with analysis and mental interpretation." In vocational selection in industry this has especially been the case. Statistics and the statistical point of view have dominated. It is the opinion of the writer that in the cause of greater scientific accuracy in vocational selection in industry the statistical point of view must be supplemented by a clinical point of view. It must be recognized that the competency of the applicant for a great many jobs in industry, perhaps even for a majority of them, cannot be observed from an objective score any more than the ability of a child to profit from one or another kind of educational treatment can be observed from such a score. There is no reason for suspecting that the capacity of an individual motorman to avoid accidents, or of a printer's apprentice to profit from instruction in this trade can be expressed in an objective score, as easily interpreted by a minor clerk as by a trained psychologist, than for suspecting that the mental status of a child is revealed in the I.Q. which can be obtained by any teacher who owns a copy of Terman's Condensed Guide and a set of testing material. The one problem is as complicated as the other; the objective score in one case has in it as many elements of error as in the other, and an adequate diagnosis in both involves interpretation by a trained psychologist based on observation of performance and a consideration of related data.

---

[6] Id., p. 23.

2. Another assumption which has led to the domination of the statistical point of view in vocational selection is that industry is not ready to pay the cost of a more refined and more detailed method. This assumption is perhaps justified in fact. The strings of industry's purse are sensitive, and it is probably true that the envisagement of a too ambitious program of vocational selection, even if it were endowed with more of the quality of scientific accuracy, might well result in the absolute closure of the purse, and the stoppage of the slight dribble of coins which have been carefully counted out for psychological work in industry. This is a real danger, but it hardly justifies the easier and the less scientific course. Education of industrial managers to the problems and needs of adequate research and the encouragement of endowed research are the ways of meeting this contingency. There may be a delay in the development of the application of psychology in industry, a diminution in the quantity of work, but this is preferable to the defense of a position which will prove ultimately to be untenable.

Industry should be told quite frankly that there are some jobs for which workers can be satisfactorily selected by means of tests graded and statistically evaluated by minor clerks in the employment office. The specific jobs for which this holds true are determined in part by the duties of the job, the universality and what might be termed the superficiality of the qualities needed for success in the job, and by certain economic factors such as cost of labor turn-over, the relative turn-over of the total group employed in this type of work, the time and cost of training, etc. But industry must also be told with equal frankness that there are a great many jobs for which workers can not be adequately selected by tests administered and interpreted by employment clerks. It must be told that the selection of applicants for such jobs involves an examination by a trained psychologist who depends upon his scientific knowledge of human behavior as well as upon the test results. Industry must be told that the immediate cost of such a procedure may be great, but that ultimate economy is implied in the acceptance of such a procedure.

3. A third assumption, implied rather than stated, an assumption which is perhaps not recognized by a great many workers in this field, is that of the relative unimportance of the individual worker in industry. A selective process which is satisfied with the probable adequate selection of a group of workers, rather than with the proper placement of the individual worker (and such is the case n selection based upon objective scores alone) fails to give adequate consideration to the well-being and interest of the individual worker, an interest which in comparison with the point of view of European psychologists, has been shamefully neglected by the American psychologist.

Other assumptions can be stated, but they can not be presented in a paper of this length.

A discussion such as this of the clinical viewpoint in industrial selection involves a reconsideration of the place of the psychologist in industry. The acceptance of the statistical viewpoint makes place in industry for what has been called a psychotechnician, a worker with a minimum training in psychology and a maximum training in statistics. He is a technician in the fullest sense of the word, well trained in method, but not in theory. He understands the various steps to be taken in standardizing a test and in statistically treating the results for the revelation of significant scores. In many cases he need know little about mental tests, since statistical treatment will reveal the significant ones, although a little information about tests may be helpful in providing shortcuts in the determination of the significant ones. Of clinical observation and interpretation of performance he need know nothing at all, and generally knows nothing. He is a consultant in industry, supervising the construction and statistical evaluation of tests and the administration of such tests by a polite, but not necessarily intelligent, clerk.

If the clinical point of view is to affect vocational selection the worker must be more than a psycho-technician. He must be a psychologist trained in psychological theory and methods as well as in statistics. He must be more than a consultant; he must be a co-worker in the personnel department, imbued through experience and study with an understanding of the selection problems of the particular plant and with the problems of the job for which he is testing. He must function in the examination of the applicant, applying his experience in the observation and analysis of performance in deciding whether a given applicant should be employed. He becomes in a sense a sort of superior employment officer, weighing the test score, his observation and analysis of performance with other essential data about the applicant in determining whether or not the applicant is competent for the job. His judgment is a diagnosis, as that of a physician, based upon a consideration of all the data affecting success or failure on the job.

These are the two extreme positions. There may be all sorts of compromises, but the necessity of the active functioning of the psychologist in the employment of the individual worker is a necessary concomitant to the adoption of the clinical viewpoint in vocational selection.

This does not mean that there remains in industry no place for what has been called the psycho-technician in industry. He has exactly the same place in industry as the laboratory technician has in medicine. He is the assistant to a trained examiner, providing such objective quantitative data as he can for interpretation by the trained examiner. In certain cases, as has been indicated, he may even be the sole representative of psychology in industry, functioning in the standardization of tests and in the determination of a set of statistical values useful in employing a particular test in selection. Such a job a psycho-technician can very well do, but his services should be provided with the understanding that industry is receiving not psychological services, but technical services; that it is using in employment not necessarily psychology but statistics slightly diluted with a smattering of psychology. Such services should not be sold under false pretenses, as has often been the case.

This paper is limited to the selection of workers for a given job. The neglect of the clinical viewpoint is equally apparent in the use of tests in vocational guidance. In this field, where a correct reading of the mental abilities and defects of the individual is perhaps more important than in industry, the need for the clinical viewpoint becomes even more urgent. It is perhaps impossible to overestimate, both from the point of view of the individual studied and the progress of applied psychology, the harmful effect, for example, of such a logical development of the statistical point of view in guidance as the program for long-distance guidance by correspondence, an advertising prospectus of which has recently been widely distributed by a well known publisher of psychological tests. It is impossible to discuss this aspect of the application of the clinical viewpoint in vocational selection within the limits of this paper, but the writer takes the opportunity to suggest the need for a re-evaluation of testing projects in vocational guidance in the light of the point of view presented herein.

ANALYZING PERSONALITY

Part I

GEORGE C. BRANDENBURG

Purdue University

How often one hears it said that Mr. A or Mr. B is a very bright man, but that he has no personality. Sometimes it is even said that such men as scientists and intellectual persons as a class are characterized to a certain extent by poor personalities. And occasionally one hears the converse, that Mr. So-and-so has a fine personality, but that there is really little to him, that he lacks depth. Promoters, reformers, and salesmen are not infrequently so described. The prominent rôle which this term personality plays in our every-day affairs is further indicated by the fact that in practically all recommendation blanks there are numerous questions concerning the applicant's personal qualities. Does he have a good personality? Is he tactful in dealing with other people? Is he forceful? Has he a positive manner? Does he coöperate readily with others? Is he industrious? Scores of such terms as occur in these questions are in common use in the literature of employment agencies and are indicative of the interest which employers generally have in other phases of the individual's make-up aside from his immediate ability to do the work in question.

What does the term personality really mean? What elements enter into it and how are they related to one another? How is personality related to intelligence? What does it have to do with success in various fields of activity? These are some of the questions which must have occurred to the reader time and again and they are likewise questions which must frequently occur to any thoughtful student of human affairs.

We use the term as if its meaning were so well understood that it required no definition, but such is far from the case. Furthermore, we frequently in our discussions use such terms as personal and character traits as if they were synonymous and could be used interchangeably. Again this is not quite true. If we examine psychological literature we find that the term, personality, has a wide variety of meanings, depending on the viewpoint of the writer. According to Warren it is "the entire mental organization of a human being at any stage of his development." Woodworth regards personality as including one's physical, temperamental, instinctive, and intellectual qualities. Watson, thinking of personality entirely from the standpoint of a behaviorist, says, "Let us mean by the term personality an individual's total assets (actual and potential) on the reaction side. By assets we mean, first, the total mass of organized habits; the socialized and regulated instincts; the socialized and tempered emotions; and the combinations and interrelations among these; and second, high coefficients both of plasticity (capability of new habit formation or altering of old) and of retention (readiness of implanted habits to function after disuse)." A most interesting and extreme view of the meaning of personality is held by those few writers who believe that it is explained wholly in terms of the activities of the ductless or endocrine glands. Of course, these investigators are more interested in showing that the glands are the basis of personality and especially of its variations, than in any definition of the term.

As used by the writer the term is understood to indicate a composite of an individual's typical reactions, physical, intellectual, emotional, to his environment, together with his various physical characteristics which constitute what we call his general appearance. Some writers go still further in this direction and assert that the term represents the sum total of the individual's reactions during his existence as a unified organism. It seems quite useless, however, and impractical to make the term so all inclusive as this; for it is quite obvious that in our actual perception of the personality of another we are limited to a frac-tional part of his total behavior. In our ordinary judgments of the personality of an individual we make use of only a few cross sections of his behavior and physical characteristics. Naturally we realize that such judgments are inadequate, but in any event they are what we use in practical every day affairs and must no doubt continue to use.

A little reflection, then, concerning the meaning of one's personality makes it clear that it depends upon both his inherited qualities and the environmental conditions which have surrounded his existence. We shall not be particularly concerned with the relative influences of these two factors in the development of personality. It should be noted, however, that one's environment has somewhat more to do with his personality than with his character; while heredity apparently plays the more important part in the matter of character.

PERSONALITY AND CHARACTER

The term character is generally used to indicate a somewhat restricted part of a person's individuality and commonly refers to those emotional and instinctive qualities which are largely innate. In any case it has little reference to the physical characteristics or the appearance. Such qualities as honesty, reliability, sincerity, and moral uprightness come to one's mind in connection with the term character. In general, then, the term personality is somewhat more objective, more akin to the term reputation and has reference to the manner in which the individual's behavior and appearance appeal to his friends and associates; it is more dependent on social judgment while character is more subjective and less susceptible to social evaluation. The term personality is associated with such qualities as enthusiasm, forcefulness, address, tact, and physical appearance in general. But there is much overlapping in the terms character and personality so that it is quite impossible to make any clear-cut distinction between them. When a sufficient number of critical studies have been made to form an adequate basis, it will, no doubt, be found desirable to set up fairly specific limitations of these terms. In any study of personality

TABLE I

Showing the intercorrelations of twenty-three personal traits

	ACCURACY	ENTHUSIASM	AGGRESSIVENESS	SELF-RELIANCE	MEMORY	POPULARITY	MOTOR ABILITY	TACT	GENERAL ABILITY	RELIABILITY	COÖPERATION	REASONING ABILITY	GENERAL INFORMATION	ORIGINALITY	SYMPATHY	SPEED IN WORK	SOCIAL AND CIVIC INTERESTS	ADDRESS	SINCERITY	INDUSTRY	NEATNESS	APPRECIATION OF HUMOR	MORAL HABITS	TOTAL PERSONALITY
Accuracy		72	86	72	83	68	67	62	70	66	72	71	58	76	64	78	72	48	33	60	33	35	36	87
Enthusiasm	72		78	71	72	76	64	81	80	66	67	65	85	66	58	71	61	50	38	53	35	28	10	86
Aggressiveness	86	78		61	72	73	67	69	53	69	52	60	53	69	57	78	62	49	25	48	45	48	30	85
Self-reliance	72	71	61		70	59	71	66	80	83	76	80	66	52	45	73	34	20	77	39	18	−02	12	85
Memory	83	72	72	70		58	64	62	47	73	52	61	50	69	66	66	47	34	54	69	27	35	41	85
Popularity	68	76	73	59	58		75	90	76	72	80	61	70	61	84	44	67	61	43	26	56	46	02	84
Motor ability	67	64	67	71	64	75		62	76	53	60	76	54	58	58	56	40	52	61	26	22	26	16	81
Tact	62	81	69	66	62	90	62		79	64	80	56	72	52	67	61	57	53	36	25	54	44	13	80
General ability	70	80	53	80	47	76	76	79		72	83	85	68	56	64	61	44	49	50	29	41	25	−24	80
Reliability	66	66	69	83	73	72	53	64	72		71	70	56	47	62	68	46	27	81	25	15	12	19	79
Coöperation	72	67	52	76	52	80	60	80	83	71		73	64	60	67	40	35	33	53	18	33	18	−27	78
Reasoning ability	71	65	60	80	61	61	76	56	85	70	73		69	57	41	66	32	47	62	26	14	−15	−18	77
General information	58	85	53	66	50	70	54	72	68	56	64	69		54	48	72	48	41	47	68	32	−08	20	76
Originality	76	66	69	52	69	61	58	52	56	47	60	57	54		71	53	70	72	45	54	12	30	13	75
Sympathy	64	58	57	45	66	84	58	67	64	62	67	41	48	71		28	68	54	36	33	19	52	13	73

Speed in work	78	71	78	72	66	44	56	61	61	68	40	66	72	53	28		48	26	52	47	32	05	26	*71*
Social and civic interests	72	61	62	34	47	67	40	57	44	46	35	32	48	70	68	48		69	−03	29	45	47	20	*61*
Address	48	50	49	20	34	61	52	53	49	27	33	47	41	72	54	26	69		−17	13	17	45	−04	*55*
Sincerity	33	38	25	77	54	43	61	36	50	81	53	62	47	45	36	52	−03	−17		26	12	−20	33	*54*
Industry	60	53	48	39	69	26	26	25	29	25	18	26	68	54	33	47	29	13	26		−17	12	60	*52*
Neatness	33	35	45	18	27	56	22	54	41	15	33	14	32	12	19	32	45	17	12	−17		−02	−22	*48*
Appreciation of humor	35	28	48	−02	35	46	26	44	25	12	18	−15	−08	30	52	05	47	45	−20	12	−02		36	*32*
Moral habits	36	10	30	12	41	02	16	13	−24	19	−27	−18	20	13	13	26	20	−04	33	60	−22	36		*17*

then, it is obvious that we shall be dealing with such traits as enthusiasm, aggressiveness, and neatness in appearance which are clearly qualities of personality. But since personality includes character, it will also be found desirable to include such outstanding qualities of character as reliability, sincerity, and moral character. At present we know comparatively little concerning the tendencies of such personal and character traits as those just mentioned to be mutually attractive, antagonistic, or neutral in one's make-up.

What likelihood is there, for example, that the intelligent person will be good, careful, tactful, coöperative, or reliable? What are the chances that the enthusiastic, and optimistic man will be sincere, sympathetic, or broadminded? How far may we expect to find the individual with an outstanding personality possessing such traits as common sense, self-reliance, and originality? To what extent is a student's grade determined by his personality? How far is a man's income dependent upon his personality?

AN EXPERIMENTAL STUDY

Group I

In an endeavor to throw some light on such questions as those just suggested, twenty-nine senior engineering students in a vocational psychology class rated each other with respect to twenty-three personal and character traits. At the close of the rating process and without any knowledge of its results, the class was given the Army Alpha Intelligence test. The traits used in the rating experiment were selected from a list of some one hundred fifty traits assembled from various rating and personnel systems. The selection was made through a process of preferential voting by the students so that those selected were judged to be the most important ones for a good personality. These traits are: Accuracy in work, enthusiasm, aggressiveness, self-reliance, memory, popularity, motor ability, tact, general ability, reliability, coöperativeness, reasoning ability, general information, originality, sympathy, speed in work, social and civic interest, address, sincerity, industry, neatness in appearance, appreciation of humor, and moral habits. Some of these traits, for example, general information, can hardly be called personal traits since they represent acquisitions of knowledge and the like rather than fixed tendencies or qualities. No attempt has been made to avoid this apparent inconsistency or to reduce the list to simple, clearly definable qualities. This is probably a weakness in the experiment. The chief justification of the traits included is that they are frequently used in rating and personnel work, are well understood, and are therefore assumed to be significant.

After the traits had been selected each member of the class agreed to rate each of the other members by assigning him a rank in each trait which would represent his standing in that particular trait with reference to the other members of the group. The rating was all done in the classroom and consumed on the average about two and one-half hours time. There was no conferring among students after the rating began and all the results were held by the instructor as confidential. After the rating sheets had been handed to the instructor the estimates were all summarized so that each student received a final ranking in the group in each trait. This final rank thus represents the pooled judgments of twenty-eight of his associates concerning his relative merit in each of the qualities. It was then possible to assign each student a final rank not only in each trait but in the total of all the traits or in personality as a whole. For the purposes of the experiment we assume that this is a correct and true measure of each individual's personality with reference to this particular group.[1]

---

[1] An individual's rank in any trait is thus the result of the combined judgments of twenty-eight of his classmates and since he was rated in twenty-three different traits, his total personality rank, that is his final rank, is based on a combined total of 644 separate judgments. It should be recalled also that twenty-seven of these twenty-nine men entered the Engineering School of Purdue University in September, 1915, and graduated in June, 1919. Two of the men entered in September, 1916, and finished in June, 1919. Thus, with the two exceptions, the men were in school together for four years. Just how intimately

As has been pointed out elsewhere there is no way of determining how far these estimates really represent valid measures of the various personal qualities. We can determine, however, by approved statistical procedure how reliable or consistent they are; that is, how far the same ratings would be again obtained under the same conditions. We have computed the coefficients of reliability for ratings made in other groups using a somewhat different array of traits.[2] We shall refer to the details of this matter a little later. It may be noted in passing that the reliability, while varying somewhat for different traits, is generally found to be high where the number of persons rating a given trait is as large as it is in this case.

Table 1 shows the intercorrelation among the twenty-three traits as well as the correlation between each and the total personality as indicated by the final rank of the individual in the various qualities. The traits are arranged in the table in the order of size of the coefficient of correlation with total personality. Referring to the table we may then conclude that the most important of these twenty-three traits as factors in the make-up of the individual's personality is accuracy in work, with a correlation of 0.87; the second most important is enthusiasm 0.86; while the least important so far as personality is concerned is moral habits, with the negligible correlation of 0.17.[3] On first thought it may seem somewhat surprising

---

the members of any group of students know one another cannot well be determined. There is much reason to believe that the men who coöperated in this rating experiment were sufficiently well acquainted to give their pooled judgments of one another a considerable degree of validity. Of course, their judgments of any individual do not agree; such judgments never do, and no degree of familiarity would produce complete agreement. In nearly every case an individual is found to be rated in a given trait all the way from the very best to the very poorest, thus indicating the unreliability of a single individual's judgment of another. This, however, is in itself no proof that the combined judgments are not highly valid.

[2] The reason the reliability was not calculated for the ratings in this particular group was that as soon as the estimates had been summarized, the original rating sheets were destroyed.

[3] In making the assumption that a given trait is relatively impor-

that moral habits as a quality is shown to be practically irrelevant in the matter of personality so far as students are concerned. Two reasons for this have occurred to the author: first, there is the characteristic reaction of the average student to the display of moral habits by another student. In spite of the instructor's warning to the contrary these students were probably inclined to think of moral habits as regular attendance at Sunday School and church, abstinence from the use of tobacco, non-use of profanity, etc. Naturally we should expect little correlation between moral habits, in such a sense, and excellence of personality. Second, it is quite possible that moral habits in the best and broadest sense of the term is not so much of a factor in a good personality as it is sometimes supposed to be. In any case it must be remembered that it is a positive factor even though it has little weight; in other words, moral habits do contribute slightly to, rather than detract from a good personality. In later experiments it has been found that the substitution of the term moral character defined as uprightness and honorableness, for moral habits increases the correlation

tant in the constitution of personality according to its index of correlation, I am fully aware of the existence of an error which runs throughout any such series of inter-correlations. This error which is always present in the correlation of a series of measures or ratings with a second series in which the first is included, naturally tends to raise all the correlations somewhat. However, with the fairly large number of traits used in this experiment, the influence of any given trait on the composite is not likely to be sufficient to affect the correlations very materially. In any case it is unlikely that the relative rankings of the various traits vary much from what they would be if the two series correlated were entirely discrete. Furthermore, since the purpose is to discover suggestive trends in the personality structure rather than exact relationships, it is obvious that the presence of this error does not justify serious criticism of the method employed.

The failure to introduce further refinement of statistical procedure is also explained by the fact that we are interested in the general nature of personality. It is quite likely that the employment of refined and complicated statistical procedure in dealing with measurements of the kind with which we are here concerned would tend to confuse rather than to clarify the situation.

with personality in a considerable degree. One is rather disappointed also that such a desirable quality as industry should not receive a more cordial welcome in this group of attractive personal qualities. We may infer that while industriousness is a very valuable asset from some standpoints, it does not tend greatly to give one an outstanding personality.

TABLE 2

Agreeable traits

	NUMBER OF VOTES
1. Honesty, fairness in grading	59
2. Courtesy and consideration	53
3. Liberal and progressive attitude	47
4. Cheerfulness	46
5. Clearness	35
6. Punctuality	18
7. Generosity	17
8. Cleanliness and neatness	15
9. Morality	7
10. Intellectuality	6

TABLE 3

Disagreeable traits

	NUMBER OF VOTES
1. Lack of courtesy in the classroom	88
2. Unfair practices in grading, etc.	69
3. Disagreeable qualities of disposition	42
4. Egotism	30
5. Irritating personal habits and peculiarities	29
6. Lack of cordiality	20
7. Lack of punctuality	20
8. Lack of dependability	18
9. Lack of public spirit	15
10. Untidiness and carelessness in dress	13

An examination of the table reveals some interesting tendencies in the interrelationships of the twenty-three traits of the individual's make-up. Perhaps the outstanding fact is the tendency of all except the last three traits to correlate to a fair or a considerable degree with each other as well as with total personality. This bears out generally the results of other studies in this respect and warrants the conclusion that desirable traits of character tend to accompany one another.

Assuming, then, as we have for the purposes of our study that these twenty-three traits make up one's total personality and that the combined estimates are reliable, we may say that the excellence of an individual's personality is dependent on the various traits roughly in the degree indicated by the following classification:

Largely by:
- Accuracy in work
- Enthusiasm
- Aggressiveness
- Self-reliance
- Memory
- Popularity
- Motor ability
- Tact
- General ability

Somewhat by:
- Reliability
- Coöperation
- Reasoning ability
- General information
- Originality
- Sympathy
- Speed in work
- Social and civic interest
- Address
- Sincerity
- Industry

Slightly by:
- Neatness
- Appreciation of humor
- Moral habits

In a similar manner the relationship of any given trait to each of the others may be shown. Popularity, for example, is determined by other qualities in this group thus:

Largely by:
- Tact
- Sympathy
- Coöperation

Somewhat by:
- General ability
- Enthusiasm
- Motor ability
- Aggressiveness
- Reliability
- Accuracy in work
- Social and civic interest
- Address
- Originality
- Reasoning ability
- Self-reliance
- Memory
- Neatness

Slightly by:
- Appreciation of humor
- Speed in work
- Sincerity
- Industry

Not at all by:
- Moral habits

Of course, such classifications of relationships are only suggestive; and it must also constantly be borne in mind that we are here dealing with a very select group; that is, these students had all practically completed four years of training in the same engineering school, were all graduates of public high schools, and had all come from somewhat similar home environments. Whether the trends in relationships between personal traits which we have found to exist in such a group would be found in a more representative population containing all grades and types of individuals can only be determined by further studies.

As a further check on the question of how various traits are related to popularity, each of forty seniors was asked to submit a list of ten qualities which he found most attractive and agreeable in his instructors; a few days later he was requested to submit another list of ten qualities which he found most disagreeable in instructors. A total of one hundred five agreeable and ninety disagreeable traits appeared in their submitted lists. Related traits were combined and the lists then summarized under ten headings for the agreeable traits and ten for the disagreeable. Tables 2 and 3 in which the various classes are arranged in order of their importance according to the number of times they were mentioned show the results.

These tables show the same general arrangement of qualities with reference to the popularity of an instructor as does the previous list with relation to popularity in general. It is apparent from both studies that such qualities as intellectuality, moral habits, industry (not mentioned in the latter lists), and neatness have little to do with one's popularity. The old adage "Virtue is its own reward" seems entirely applicable in this case. Consideration for and sympathy with others, courtesy, a coöperative spirit, enthusiasm, and cheerfulness, fairness and reliability are the traits which are most likely to make one a favorite with his associates. One would suppose that intellectuality in an instructor would surely be one of the first qualities a student would look for and admire; but after all students are only human and apparently crave first the same humane qualities in their instructor that they demand in their classmates.

Group II

One year after the study just described a second group of twenty-five senior engineering students rated each other following the same technique as group I, with the exception that nine traits were used instead of twenty-three. The traits used in the second group are by definition made somewhat more inclusive than were those formerly used. They are: common sense, judgment, or tact; trustworthiness, reliability, or dependability; executive ability, planning and directing; general intelligence, mental caliber, capacity to learn; resourcefulness, originality, or initiative; aggressiveness, energy, or industry; broadmindedness, tolerance, open-mindedness; confidence or self-reliance; moral character, uprightness, honorableness.

TABLE 4

	COMMON SENSE	ORIGINALITY	AGGRESSIVENESS	EXECUTIVE ABILITY	BROAD-MINDEDNESS	SELF-RELIANCE	INTELLIGENCE	RELIABILITY	MORAL CHARACTER	TOTAL
Common sense		93	88	90	83	84	88	84	65	*95*
Originality	93		94	92	83	92	87	72	52	*95*
Aggressiveness	88	94		89	83	92	86	74	56	*94*
Executive ability	90	92	89		79	88	85	85	51	*92*
Broad-mindedness	83	83	83	79		83	73	83	57	*90*
Self-reliance	84	92	92	88	83		88	70	49	*89*
Intelligence	88	87	86	85	73	88		70	50	*88*
Reliability	84	72	74	85	83	70	70		74	*85*
Moral character	65	52	56	51	57	49	50	74		*66*

Table 4 shows the intercorrelation among these traits together with the correlation of each with total personality. It will be noted that the correlations are uniformly somewhat higher than those in the first group, but that the general trend is similar. Those qualities which in their display compel attention, the qualities of enthusiasm, aggressiveness, and originality when combined with common sense, tact, and accuracy in work, are most prominent; while moral character is of lesser prominence in the outstanding personality. The fact that moral character shows a fairly good correlation with total personality in this group whereas moral habits showed a negligible one in the first group is probably an indication of the difference between the interpretation of the two terms by the raters. Moral character is presumably a more acceptable term as a personal quality to the student than is the term moral habits.

While it is entirely likely that there is a considerable tendency for such desirable traits as we are here considering to accompany each other in one's personality, it is very doubtful that they are interrelated in such a high degree. These high indices are more likely produced to some extent by what is known among psychologists as the "halo" influence; that is, the tendency for an individual to judge the specific details of another's character according to his general impression of him. This is presumably more likely to be the case where the rating is all done at one sitting rather than in several periods separated from each other by considerable intervals of time.

TABLE 5

Showing the reliability coefficients of the ratings in nine personal qualities in a group of twenty-five students

QUALITIES	RELIABILITY COEFFICIENTS
General intelligence	0.87
Aggressiveness	0.85
Executive ability	0.75
Originality	0.74
Reliability	0.73
Common sense	0.67
Moral character	0.64
Self-reliance	0.63
Broadmindedness	0.58

The reliability of the judgments in this group has been calculated by the well-known method of correlating the combined ratings of one half of the group with the combined ratings of the other half. The table (table 5) showing the coefficients of reliability for the different traits suggests that our results are entirely satisfactory in this respect.

PERSONALITY AND GENERAL INTELLIGENCE

Common observation would lead one to expect that a good personality would involve at least a fair degree of general intelligence and that a high type of general intelligence would be more likely to be found in a person with a good, than in one with a poor personality. It will be recalled that general intelligence as estimated by one's associates, and personality are very closely related, the coefficient being 0.88; but the correlation in this case is, no doubt, affected favorably by the fact that general intelligence was estimated in connection with the rating of other personal traits. It has already been suggested that the "halo" influence tends to increase the correlations between desirable traits under such conditions. Table 6 showing the correlations between an objective measure of general intelligence, the Army Alpha Test, and personality in these two groups undoubtedly gives us a more accurate notion of the tendency of these two qualities to accompany each other.

TABLE 6

Showing the correlations between personality, general intelligence, and school grades for groups I and II

GROUP	CASES	PERSONALITY AND INTELLIGENCE	PERSONALITY AND SCHOOL GRADES	INTELLIGENCE AND SCHOOL GRADES
I	29	0.51	0.61	0.56
II	25	0.34	0.57	0.45

SCHOOL GRADES

Reference to table 6 indicates that there is slightly closer relationship between the student's personality and his grades than between his intelligence (measured objectively, not estimated) and his grades. Does this mean that those students who are sometimes accused of deliberately trying to "work" the instructor for a good grade succeed? To ascertain whether this is the case or not we have correlated the various personal traits with school grades for group I. This will indicate which traits are most effective in the production of high grades assuming that the traits do have a causal relationship to the grades.

If we take those traits showing correlation of 0.50 or better as being of considerable importance, those between 0.30 and 0.49 as having a moderate influence and those below 0.30 as being negligible or of slight importance in the matter of grades, we may then say that, in so far as grades are influenced at all by personal traits, the relative influence of such traits is substantially as follows:

Considerable influence		*Moderate influence*		*Negligible or slight influence*	
Reasoning ability	0.76	Accuracy	0.45	Popularity	0.20
Speed in work	0.73	Originality	0.43	Neatness	0.10
Motor ability	0.59	General ability	0.43	Coöperativeness	0.10
Self-reliance	0.59	Aggressiveness	0 42	Sympathy	0 13
General information	0.57	Reliability	0.41	Appreciation of humor	0.00
Industry	0.50	Enthusiasm	0.34		
Sincerity	0.50	Moral habits	0.33		
		Memory	0.32		
		Address	0.32		
		Social and civic interest	0.31		
		Tact	0.31		

This array of correlations presents splendid evidence that the instructor is not very susceptible in his grading to the diplomatic endeavors of students. The student may laugh ever so heartily at the instructor's stale jokes, he may be ever so sympathetic with his pet theories, and he may present an immaculate appearance in class; but apparently all this will avail him naught so far as grades are concerned if he hath not reasoning ability, mental agility, self-reliance, a good fund of general information, and industrious spirit, and some honesty of purpose. There are, no doubt, instructors who do permit irrelevant personal factors to enter into their grading, but the probabilities seem to be that such instructors are rare as compared with those who disregard all those factors which do not make for substantial scholastic attainment.

How far these results would be valid for other types of schools is problematical. Engineering subjects are of such a nature generally hat achievement in them can be graded rather objectively and accurately. It is likely that the student finds it much easier to "work" his instructor in subjects which involve a good deal of philosophizing and theorizing than in those which require mathematical precision. The writer is inclined to think that such attempts to influence instructors are much less frequent than they are supposed to be and are confined largely to the inferior students.

Having made our analysis of these two groups we shall next attempt to discover how the performances of the men in real life correspond with our various estimates and measures; for after all, the manner in which one stands the tests in the field, in the shop, and in the office will always be regarded as the real criterion of his ability. The measurement, however, of an individual's success under the infinitely complex conditions of business and industrial life presents well nigh unsurmountable difficulties. What progress we have been able to make in this direction will be described in the following section.

(*To be continued*)

GROUP DIFFERENCES IN INTELLIGENCE TESTS

The Relative Difficulty of Types of Questions

GRACE E. MANSON

The University of Michigan

INTRODUCTION

Group intelligence tests can be made to yield more information about individuals and groups than is derived from the final scores. As psychologists, we have overlooked a good deal of valuable information because we have not paused to scrutinize the examinee's reactions to the test situation in detail.

Of course it is valuable to know that a group of expert stenographers makes, on the average, about the same intelligence test score as a group of civil engineers; or that nurses and bookkeepers in the Army were found to be of about the same level and range of intelligence. But the tests can tell us more than this. Indeed, we are more interested in finding out how the nurses differ from the bookkeepers than in finding out the respects in which they are alike. This paper reports some differences—other than differences in intelligence test scores—between two groups of graduate students, a group of industrial engineering students, a group of mechanical engineering students, a group of men engaged in fitting themselves to be salesmen of life insurance, and other groups consisting of minor executives in a manufacturing company, high school students and policemen. The members of each of these groups have in common some such factor as academic training, vocational interest or occupational pursuit. We have raised the question whether their performance in group intelligence tests cannot be made to yield detailed information concerning specific abilities and disabilities. Are there significant group differences in the relative difficulty of the several types of tasks comprising the intelligence examination?

The answers to such questions would be of value to both the theoretical psychologists and the vocationalists. The former are interested in knowing more concerning the general laws governing interests, motives, attitudes, effects of training, ability to profit by experience, and a host of allied problems; all of which laws are discovered and verified from observing many selected groups in many controlled situations. The vocationalists have to be more specific, and are anxious to find means of ascertaining the interests, motives, attitudes, training, capabilities, etc., of their clients, as well as the methods of putting these forces to the best use when once analyzed. For example, on the basis of final scores, the army psychologists,[1] also Fryer[2] and Burr,[3] have been able to determine the minimum and the maximum and the range of intelligence test scores required for many occupations. But they have also found that the range of test scores at any occupational level is so wide that it includes scores common to many other occupations. It, therefore, becomes necessary to both subject the intelligence examination to more detailed analysis and to devise means for supplementing the intelligence tests by measures of personality and character.

We have limited ourselves to but one phase of the former task, namely, to find the relative difficulty of the several types of questions for groups of different interests and educational status.

---

[1] Army Mental Tests, Methods, Typical Results and Practical Applications, November 22, 1918, Washington, D. C

[2] Fryer, Douglas, Occupational-Intelligence Standards, School and Society, 1922, Vol. 15, pp. 273–277.

[3] Burr, Emily T., Minimum Intellectual Levels of Accomplishment in Industry, Journal of Personnel Research, 1924, Vol. 3, No. 6, pp. 207–212.

PREVIOUS INVESTIGATIONS

Cobb and Yerkes[4] have been interested in the same problem, with special reference to army officers. They compared the median percentile score made by officers in each branch of the service with the standard score made by all branches taken together (considered the 50 percentile) and plotted psychographs showing the success attained by officers in each group. Their results lead them to conclude that "the psychograph (curve representing measurements for eight types of test which constitute army group-examination) for medical officers differs strikingly from that for artillery officers or from that for engineers. The psychographs of the several special medical groups have a very obvious family resemblance."

Bingham and Davis[5] have carried the investigation into the realm of industry. They studied the intelligence test records of 102 business men attending the annual statistical conference held at the Babson Institute, August 4, 1922. The personal history records of seventy-three of these men were sufficiently complete to classify them in respect to present occupation as business executives, salesmen or doubtful cases. The educational record of each of these men was also known.

The intelligence examination used was not Army Alpha but an adaptation of it known as Bureau of Personnel Research Test VI. Eighty-three per cent of these business men made scores on Test VI which, transmuted into Army Alpha scores, gave them a rating of A or B. The median scores of the three occupational groups were about the same. Examination of the personal history records showed that these men were, on the whole, superior business men. We have here a group of successful executives and salesmen of superior intelligence. In these two respects they are very much alike. What does an analysis of the scores made on the six parts of the intelligence examination show? In their conclusions, the authors state:

---

[4] Cobb, M. V., and Yerkes, R. M., Intellectual and Educational Status of the Medical Profession as Represented in the United States Army, Bulletin of the National Research Council, 1921, Vol. 1, Part 8, No. 8, pp. 458–532.

[5] Bingham, W. V., and Davis, W. T., Intelligence Test Scores and Business Success, The Journal of Applied Psychology, 1924, Vol. 8, No. 1, pp. 1–22.

> The salesmen, as a group, were more proficient upon disarranged sentence questions than were the executives. The executives, on the other hand, were more proficient upon arithmetic and number completion questions. Proficiency on the arithmetic questions shows a negative correlation with schooling beyond grammar school. Proficiency in opposites differentiates the occupational groups not at all but shows some correlation with schooling.

PROBLEM

In the present paper, the writer reports on the relative difficulty of the several types of questions on Bureau of Personnel Research Test VI, as shown by the scores made by eight groups of persons of different ages, educational status and occupational interests. It has been our purpose to see if groups having some factor in common, as education or occupational interest, would find the same types of questions difficult.

SUBJECTS

The test records of 341 persons were studied. These persons were divided into eight groups, as follows:

Group I. Retail bureau students

This group consists of 14 persons, 13 women and 1 man. All of these were college graduates, 10 of them candidates for their Master's degree at the University of Pittsburgh. All of them were interested in preparing to hold executive positions in the personnel department in retail stores.

Group II. Industries students

This group is composed of 28 seniors in the College of Industries at Carnegie Institute of Technology. They were all men in the twenties or thirties, interested in becoming executives or minor executives in the trades of printing, plumbing, building, or machine production.

Group III. Mechanical engineers

Thirty-six seniors in the school of mechanical engineering at Case School of Applied Science make up this group.

Group IV. Carnegie Life Insurance School students

This group includes 36 men who attended the Carnegie Life Insurance School. Many of them were either college graduates or else had attended college. Some of these men had already sold insurance and were sent to the school by their companies, others were taking the course preparatory to selling life insurance. All of them had sufficient interest in life insurance as a vocation to undertake the eleven weeks intensive course.

Group V. Manufacturing company office executives

This group consists of 40 members of the "office" group of a large company manufacturing heating appliances. These men were in the offices of the company in minor-executive positions such as inspectors, supervisors and assistant managers. They were not actively selling but were interested in sales, production and service.

Group VI. Policemen

This group contains 81 policemen of various grades up to lieutenant.

Group VII. Education students

This group is composed of 25 members of a course in mental measurements at the University of Michigan Summer School. In the class were both men and women. Most of them had had several years of experience in teaching or school administration.

Group VIII. High school students

Eighty-one high school sophomores, both boys and girls, make up this group.

THE INTELLIGENCE TEST USED

We have used the same intelligence examination as was used by Bingham and Davis.[6] Bureau of Personnel Research Test VI is a modified Army Alpha in which the questions on arithmetic, opposites, analogies, true-false (disarranged sentences), information and number series are arranged in a seemingly irregular order. The examination is of the spiral omnibus form, each question being slightly more difficult than the question of the same type which preceded it. Correct answers are not weighted but those requiring more time, as arithmetic, number series and true-false, are fewer in number: 20, 20, and 24 as against 40, 40 and 40 for opposites, analogies and information. True-false and opposites, following the army scheme, are scored R-W. Omissions on all tests count as errors but are not doubly penalized in the true-false and opposites tests.

---

[6] Ibid., pp. 8–9. Also described in Bingham, W. V., Some Group Tests, Psychological Bulletin, February, 1920, p. 57.

Test VI has been standardized and is published in two forms, Form A and Form B. We have used Form A. The usual time allowed is four minutes for reading the directions and fifteen minutes for the test proper. Fifteen minutes has been our time limit in all groups. However, Hansen and Ream[7] have found that the time allowed did not materially change the standing of persons within a group. Total score on the complete test correlated 0.88 with score at end of 5 minutes, 0.92 with score at end of 10 minutes, 0.96 with score at end of 15 minutes, and 0.97 with score at end of 20 minutes.

METHOD

In determining the relative difficulty of the six types of questions, account has been taken, (1) of the speed factor as shown in the number of problems covered, (2) of the inability factor as shown by the number of examples omitted and in error.

The statistical measure used has, in general, been the average percentage of the group. The average in each case has meant the arithmetic mean. The procedure followed in handling the test records was as here described. Look first at number of the last question attempted. From the key, find out how many arithmetic, analogies, opposites, etc., have been included. Count how many errors have been made on each of the six types of question. Find the number of omissions on each type of question. Figure the per cent of omissions on each type of question by dividing the number of questions of that type included up through the number of the last question the subject attempted, into the number of that type omitted. Figure per cent of error on each type of question by dividing number included (same denominator as for omissions) into number of errors for that type. In the case of true-false and opposites, multiply the number of errors by two before dividing.

---

[7] Hansen, C. F., and Ream, M. J., The Predictive Value of Short Intelligence Tests, Journal of Applied Psychology, June, 1921, Vol. 5, pp. 184–186.

TABLE 1

Average scores on Test VI

GROUP	AVERAGE SCORE	
I	124.2	Retail Bureau
II	111.6	Carnegie Industries
III	112.5	Case Mechanical Engineers
IV	105.0	Carnegie Life Insurance Salesmen
V	94.8	Manufacturing Company Minor-executives
VI	38.5	Policemen
VII	121.6	Michigan Summer School 1924
VIII	75.0	High School Sophomores

There were, then, three measures for each subject, (1) number of examples covered, (2) percentage of each type of example omitted, (3) percentage of each type of example wrong. Besides this there was, for each of the eight groups, (1) the average number of examples attempted, (2) the average per cent of omissions on each type of question, and (3) the average per cent of errors on each type of question.

The final score for each subject, using the usual scoring method for Test VI, had previously been determined, as well as the average score made by each group. These average final scores are given in table 1.

Group I and Group VII have approximately the same average score. These are the Retail Bureau and the Michigan Summer Session groups. Both groups are composed of many college graduates. Groups II and III have average scores of about the same size. These are the two applied science groups, industries and mechanical engineering. Group IV is the Life Insurance group, Group V the lesser executives group. Group VIII, the high school group, is next lowest with Group VI, the policemen, as the most inferior group.

These means are in keeping with the norms found for comparable groups with Test VI. The life insurance group here is above the expected norm having a mean of 105, whereas a mean of 105 falls in the eighth decile for a group of 537 members of the Carnegie Life Insurance School. The norms for 100 college seniors place scores of 111 and 112 in the eighth decile. But since engineering students are possibly a more selected group, these means are not particularly high. Scores of 121 and 124 are in the ninth decile for college seniors which would be about where we would expect to find graduate student groups like Group I and Group VII. A mean of 75 puts the high school group in the second decile of college seniors which is about what we would anticipate. The minor-executive group with a mean of 94.8 is to be compared with a mean of 107 found for salesmanagers at a Babson conference and of 75 for salesmen belonging to the same company as these minor-executives. The policemen are much lower than the high *school students and come in the second decile of a group of* order taking salesmen, which is probably the easiest type of selling.

As a whole we should say that these groups are representative of their classes with the exception of the life insurance salesmen who have scores somewhat too high.

THE SPEED FACTOR

Since all groups were given the tests under the same time conditions, fifteen minutes, a comparison of the number of the last example attempted would give some evidence of differences in the speed factor.

The difference between the number of the last example attempted and the final score is equal to the sum of the errors and the omissions. The size of these differences follows the inverse order of the number of the last example attempted. Groups which have attempted the most examples have made the smaller number of errors and omissions. This is in keeping with previous experimental findings that speed and accuracy are highly correlated. The life insurance salesmen and the minor-executives change places in this relationship, the insurance group has made more errors and omissions than expected, the minor-executives group less.

Test VI is so arranged that it is impossible to determine the speed factor for each type of question. However, the general conclusion can easily be made that the difficulty of the questions for the group is a determining factor since groups which have covered little ground have also made many errors and omissions.

TABLE 2

Average number of last example attempted

	GROUP I	GROUP II	GROUP III	GROUP IV	GROUP V	GROUP VI	GROUP VII	GROUP VIII
Number of last example attempted............	140.5	127.5	130.2	120.6	112.4	67.5	137.8	99.2
Final score............	124.2	111.3	112.5	105.0	94.8	38.5	121.6	75.0
Difference............	16.3	16.2	17.7	21.5	17.6	29.0	16.2	24.2

RELATION OF NUMBER OF OMISSIONS TO NUMBER OF ERRORS

These facts are quite clearly brought out in the next table which shows the relation of the number of omissions to the number of errors for each group, and the relation of each of these to the number attempted (table 3).

Table 3 gives the gross averages of errors and omissions. The total number attempted is not taken into account. This relationship will be shown in percentage form in a later table (table 4). The interesting points to note are:

1. That the average number of omissions is high where the average of the last attempted is low, Groups V, VI and VIII.

2. That where the average number of last attempted is low, the ratio of errors to omissions is high. Group I is an exception to this.

TABLE 3

Average number of errors and omissions

	GROUP I	GROUP II	GROUP III	GROUP IV	GROUP V	GROUP VI	GROUP VII	GROUP VIII
Omissions	3.8	1.2	1.3	1.9	5.6	11.0	2.2	5.4
Errors	12.5	15.0	10.4	19.7	12.0	18.0	14.0	19.0
Number last attempted	140.5	127.5	130.2	126.6	112.4	67.5	137.8	99.2
Final score	124.2	111.3	112.5	105.0	94.8	38.5	121.6	74.8

TABLE 4

Average percentages of averages of errors and omissions

	GROUP I	GROUP II	GROUP III	GROUP IV	GROUP V	GROUP VI	GROUP VII	GROUP VIII
Omissions	0.027	0.009	0.009	0.015	0.05	0.105	0.016	0.054
Errors	0.081	0.117	0.126	0.156	0.107	0.266	0.101	0.191
Number last attempted	140.5	127.5	130.2	126.6	112.4	67.5	137.8	99.2

3. That groups of approximately the same speed show individual differences in their choice between omitting and making errors, Groups I and VII.

Table 4 shows these same relationships in percentage form. This per cent of error means per cent of total number covered (number of last question attempted) that are wrong. Per cent omitted means per cent of total number covered (number of last question attempted) that are omitted. Thus, account is taken of the number covered when considering the number omitted and wrong.

This table brings out very clearly several points:

1. That, on the whole, knowledge of the material governs the speed factor, which is shown by the number of examples covered, since percentages of both errors and omissions are low where speed is high. See Groups I, II, III and VI.

2. That when the examples become very difficult for the group, the number of omissions and the number of errors both increase, the ratio between the per cent of omissions and per cent of errors approaching a 1:1 condition. A 1:2 ratio is found for Groups V and VI.

3. That groups show individual differences in choosing between omissions and errors, showing that it is likely that factors of "caution"[8] and "rashness" are operative. Groups IV and V are noteworthy examples of this. It is questionable whether this urge, to be cautious or rash, becomes operative to any degree until the material reaches a certain degree of difficulty. A more detailed discussion of this relationship will be found in a later paper.[9]

RELATIVE DIFFICULTY OF THE TESTS AS SHOWN BY OMISSIONS

Another point of interest is to determine whether the high percentages of errors and omissions concentrate on the same type of questions in all groups. In the next table, table 5, is given the percentages of omissions of the total number tried for each group, arranged in the order of increasing amounts. It should be noted that this table shows relative difficulty within each group, not between groups.

If we study the concentration of omissions by comparing the five groups making the highest scores (all above 100), Groups I, II, III, IV and VII, with the three groups making the lowest score (all below 100), Groups V, VI and VIII, certain interesting points are brought out.

---

[8] Brown, Wm., A Study of the "Caution" Factor and Its Importance in Intelligence Test Performance, American Journal of Psychology, 1924, Vol. 35, No. 3, pp. 368–381.

[9] Manson, Grace E., The Significance of "Omissions" on a Group Intelligence Test.

1. The high score groups have their larger percentages of omissions in the number series, information and arithmetic tests, their small percentages of omissions on the true-false, analogics and opposites tests. The small percentage of omissions on information items in Group III is the only exception to this.

2. The groups making lower total scores are not as consistent but have certain points in common. Opposites and true-false remain in the few omissions class but analogies moves to the greater omissions class, having the largest number of omissions in Groups V and VI. Arithmetic and information remain in the more omissions class, the same as with superior subjects.

TABLE 5

Average percentages of averages of omissions on each type of question

GROUP I	GROUP II	GROUP III	GROUP IV
T. F. 0.000	Opp. 0.000	Inf. 0.001	Anal. 0.000
Anal. 0.0025	T. F. 0.0018	Opp. 0.002	T. F. 0.000
Opp. 0.005	Anal. 0.005	T. F. 0.003	Opp. 0.002
No. S. 0.038	Inf. 0.009	No. S. 0.015	Inf. 0.01
Inf. 0.052	No. S. 0.013	Anal. 0.018	No. S. 0.033
Arth. 0.09	Arth. 0.047	Arth. 0.028	Arth. 0.035
GROUP V	GROUP VI	GROUP VII	GROUP VIII
No. S. 0.012	Arth. 0.03	Anal. 0.000	T. F. 0.01
Opp. 0.024	Opp. 0.09	T. F. 0.000	Opp. 0.02
T. F. 0.027	T. F. 0.098	Opp. 0.01	No. S. 0.03
Arth. 0.044	No. S. 0.12	Arth. 0.019	Anal. 0.04
Inf. 0.05	Inf. 0.22	No. S. 0.024	Arth. 0.06
Anal. 0.111	Anal. 0.28	Inf. 0.04	Inf. 0.13

True-false and opposites are consistently among the fewer number of omissions in each of the eight groups. The order is sometimes reversed but in all groups true-false and opposites are found among the three lowest percentages of omissions. This is no doubt due to the fact that the guessing element is made very evident in the form of the questions, the 50-50 chance of being right is reacted to by the majority of subjects. But when the words are very difficult for the subjects, even this lure is ignored, for policemen omitted 9.8 and 0 per cent of the true-false and opposites questions.

On the other hand, arithmetic examples are omitted more often than other types of questions, that is, they fall among the highest three percentages for all groups except the policemen. Number series has somewhat the same tendency although not as decided. These two types of tests are the least subject to guessing and also take more writing. Thus, for persons interested in speeding up, these would be logical places to omit.

Analogies and information items offer a 1:4 guessing proposition. If realization of the chance element were operative alone, we should expect analogies and information items to stand third and fourth in the rank of omissions. But from their positions in the table, it seems likely that the difficulty of the operations had more influence. For analogies are omitted more often than any other items by the policemen and the minor-executives groups and least often by the life insurance and Michigan graduates groups. Information test percentages are irregular in position. In five out of the eight groups, they are ranked last or next to last. Difficulty therefore seems to be the stronger force.

Thus we might conclude that the percentage of omissions is controlled, (1), by the form of the question, that is, likelihood of success in guessing; (2), by the difficulty of the type of question; (3), by the time that the type of question required in solving; (4), by the operating strength of the directions "not to omit;" and (5) by other reasons particular to the group or to individuals in the group.[10]

RELATIVE DIFFICULTY OF THE TESTS AS SHOWN BY ERRORS

In this section the relative difficulty of the several types of questions will be shown by the percentages of error. The chance factor of success on true-false and opposites was eliminated by the scoring used; each error deducted two from the total score. The percentages of error are shown in table 6.

---

[10] Supporting evidence gained through introspective reports is given in the article previously referred to.

It is to be noted that opposites fall among the three items having the largest percentages of error in three of the superior groups, I, II and III, and rank third in Groups IV and VII. True-false ranks second in Groups I, II and VII, fourth and sixth in Groups III and IV. In the three low score groups, opposites rank respectively 2, 4 and 6, true-false as 6, 6 and 3. Either true-false or opposites rank among the higher percentages of error in all of the low score groups and in all but one of the high score groups. Thus, when the guessing element is penalized in the scoring, true-false and opposites are not as easy as one might expect from looking at omissions alone.

TABLE 6

Average percentages of averages of error on each type of question

GROUP I		GROUP II		GROUP III		GROUP IV	
Anal.	0.04	Anal.	0.07	Arth.	0.077	Anal.	0.11
T. F.	0.068	T. F.	0.096	No. S.	0.079	Arth.	0.12
No. S.	0.07	No. S.	0.11	Anal.	0.08	Opp.	0.134
Arth.	0.09	Arth.	0.114	T. F.	0.088	No. S.	0.15
Opp.	0.10	Opp.	0.118	Opp.	0.130	Inf.	0.18
Inf.	0.16	Inf.	0.19	Inf.	0.23	T. F.	0.228
GROUP V		**GROUP VI**		**GROUP VII**		**GROUP VIII**	
Arth.	0.066	Arth.	0.076	Anal.	0.03	Anal.	0.11
Opp.	0.09	Inf.	0.18	T. F.	0.054	No. S.	0.12
Inf.	0.10	No. S.	0.20	Opp.	0.10	T. F.	0.146
No. S.	0.12	Opp.	0.30	No. S.	0.13	Arth.	0.15
Anal.	0.12	Anal.	0.35	Arth.	0.16	Inf.	0.22
T. F.	0.134	T. F.	0.44	Inf.	0.16	Opp.	0.28

Analogies and information items are scored in the same manner for errors as for omissions. Information items are difficult for the superior groups, being the most difficult type of question for four out of five of these groups and next to the most difficult for the other one. In the case of the low score groups, the information items rank 3, 2 and 5. In relation to the other types of questions, the information questions are harder for the high score groups than for the low score groups. The reverse condition is true for analogies. Analogies rank 1, 1, 3, 1 and 1 for the high score groups, and 5, 5 and 1 for the low score groups.

The tests dealing with numbers, arithmetic problems and number series, are probably the most time consuming of all the tests. Nevertheless, arithmetic problems are the easiest kind of questions for the minor-executives and policemen, and hold an intermediate position for the high school group. In the high score groups, arithmetic problems fall in the medium percentage groups, ranking 4, 4, 1, 2 and 5. For two of the low score groups, arithmetic problems are easiest. For three of the high score groups, they are among the more difficult. Number series ranks about third with the high score groups, standing 3, 3, 2, 4 and 2. In the low score groups, number series stand 4, 4 and 1. We may then conclude that the arithmetic questions are easier than the other types of questions for the low score groups and among the harder types of questions for the high score groups. Number series are slightly easier for the high score groups than for the low score groups. On the whole, however, we may say that on the basis of errors alone, arithmetic and information are among the hardest tests for the high score groups and among the easier tests for the low score groups. The number series test takes an intermediate position. This conclusion is not without exceptions, the Case group and the high school group not following the order of their classification.

RELATIVE DIFFICULTY OF THE TESTS AS SHOWN BY BOTH ERRORS AND OMISSIONS

If we combine errors and omissions, certain very interesting contrasts stand out, as shown in table 7.

The most important point that this table brings out is that the high score groups find the word abstractions, like analogies, true-false and opposites, the easier type of questions, whereas the low score groups find them the more difficult types. Policemen and minor executives find the arithmetic, number series, and information items the easier. These are more concrete types of question, dealing very much less with the finer, more exact knowledge of words and written language.

Interesting exceptions to the general tendency are found in Group III, the mechanical engineering group, and Group VIII, the high school group. In the Case engineering group, success in opposites, an ability dependent upon word knowledge, drops below number series and arithmetic. Since mathematics is stressed in the engineering schools, rather than the cultural subjects, this condition seems logical. The other possibility is that persons interested in figures rather than niceties of words go into the engineering schools. The high school group, a mixed and undeveloped group, is irregular, neither the word nor the number problems predominating.

TABLE 7

Average percentages of averages of errors and omissions combined

GROUP I		GROUP II		GROUP III		GROUP IV	
Anal.	0.0425	Anal.	0.075	T. F.	0.091	Anal.	0.11
T. F.	0.068	T. F.	0.0978	No. S.	0.094	Opp.	0.136
Opp.	0.105	Opp.	0.118	Anal.	0.098	T. F.	0.155
No. S.	0.108	No. S.	0.123	Arth.	0.105	Arth.	0.165
Arth.	0.18	Arth.	0.161	Opp.	0.138	No. S.	0.183
Inf.	0.212	Inf.	0.199	Inf.	0.231	Inf.	0.19
GROUP V		GROUP VI		GROUP VII		GROUP VIII	
Arth.	0.110	Arth.	0.106	Anal.	0.03	No. S.	0.15
No. S.	0.132	No. S.	0.32	T. F.	0.054	Anal.	0.15
Inf.	0.15	Opp.	0.39	Opp.	0.11	T. F.	0.156
Opp.	0.158	Inf.	0.40	No. S.	0.154	Arth.	0.21
T. F.	0.161	T. F.	0.538	Arth.	0.179	Opp.	0.30
Anal.	0.245	Anal.	0.63	Inf.	0.20	Inf.	0.35

Looked at from another point of view, the school and the business groups, we find that the five school groups, I, II, III, VII and VIII are better at handling word relations than at handling arithmetic and general information items, whereas the outside groups, IV, V and VI are better at solving questions of general information and arithmetic. Group IV, the life insurance group, follows the rule of the school groups. This is likely due to the fact that it is a Life Insurance School group. The high school group, although more like the school group, has some of the characteristics of the "outside groups." The mechanical engineers have some of the characteristics of the school groups, some of the "outside groups."

As a general conclusion, we might say that the academic groups have been more successful in handling abstract word relations than in handling items of number relation and general information, and that the "outside groups" have been more successful in solving arithmetical relations and answering questions of general information than in solving problems which involve the handling of abstract word relations. This tendency is particularly strong in the case of analogies and arithmetic problems. The impression should not be gained, however, that the academic groups cannot do arithmetic problems or that the outside groups are apt at them. The mean score of Group I (high score) for arithmetic problems is, for example, 12.5 whereas for Group VI (low score) the mean score is 6.67, yet the first group has made a lower mean score on arithmetic than on any of the other types of questions and the other group has made a higher mean score on arithmetic than on any of the other questions.

Reasons for these differences in the relative difficulty of the several types of problems are more in the form of conjectures and hypotheses than in the realm of established facts. Immediate interests, which are always a governing factor, may have caused one group to work one type of question with more ease than another type. The familiarity of the content and methodology is another likely cause of differentiation.

The influence of interests makes one wonder if a change of environment would change the type of question answered most successfully. Would a group of college graduates, who had become successful business executives, excel in arithmetic and general information rather than in analogies, opposites or true-false examples? Bingham and Davis[11] have taken up this problem, considering a group of successful salesmen and executives who had taken Bureau Test VI. They report: "Proficiency in arithmetic questions shows a negative correlation with schooling beyond grammar school." "Proficiency in opposites differentiates the occupational groups not at all, but shows some correlation with schooling." There then seems to be some evidence that this greater interest and facility in the abstract holds after the college period.

SUMMARY

Are there group differences in the relative difficulty of the several types of tests comprising an intelligence examination?

To answer this question, we have studied the mental test records of 341 persons, comprising eight groups of different occupational and educational interests. The relative difficulty of the six types of questions has been determined from the number of examples attempted, the number of errors and the number of omissions. Although the number of cases included and the differences between steps in difficulty are not great, certain tendencies are sufficiently outstanding to be of some significance. The value of some of these findings is further enhanced by the fact that they verify the results of Cobb and Yerkes, and Bingham and Davis. In comparison with these investigators, our results show:

1. That seniors in mechanical engineering find opposites and information items the most difficult types of question; number series and arithmetic problems among the easier problems. The order of difficulty from easiest to hardest is: (*a*) True-false, (*b*) Number Series, (*c*) Analogies, (*d*) Arithmetic, (*e*) Opposites, (*f*) Information. Cobb and Yerkes,[12] with a very much larger group including men in all branches of en-

---

[11] Bingham, W. V., and Davis, W. T., Intelligence Test Scores and Business Success, Journal of Applied Psychology, 1924, Vol. 8, No. 1, pp. 1–22.

[12] Op. cit., p. 478.

gineering, found the order of increasing difficulty to be: (*a*) Number Series, (*b*) Analogies, (*c*) Arithmetic, (*d*) True-false, (*e*) Information, (*f*) Opposites.

2. That insurance salesmen, Group IV, find analogies, opposites and true-false questions easier than arithmetic, number series and information items. The order of increasing difficulty is as given in the preceding sentence. Bingham and Davis[13] report: "Thus, we suggest that disarranged sentences, arithmetic and number completion problems be included in tests designed to differentiate prospective salesmen from other business men, the hypothesis being that superiority in disarranged sentences, coupled with inferiority in arithmetic and number completion, is more often typical of salesmen."

3. That minor-executives, Group V, find arithmetic, number series and information problems among the easier types, with opposites, true-false and analogies as the more difficult types. Bingham and Davis[13] found "The executives, on the other hand, were more proficient upon arithmetic and number completion questions."

From the point of view of educational status, we have found:

1. That groups of graduate students, whose college courses have been in the Arts College or in Schools of Education, find the order of increasing difficulty to be as follows: (*a*) Analogies, (*b*) True-false, (*c*) Opposites, (*d*) Number Series, (*e*) Arithmetic, (*f*) Information. This is the order for Groups I and VII. Group II, which is composed of industries students, follows the same order. In this group, the academic influence seems to be stronger than the influence of the vocational interest.

2. That mechanical engineering students follow more closely the order for engineers than for graduate students specializing in the more philosophical subjects.

3. That high school students, Group VIII, find opposites difficult and number series easy, which makes them more like the mechanical engineering students than like the Arts students.

4. That, as a whole, the "in-school" groups, I, II, III, IV,

---

[13] Op. cit., p. 21.

VII and VIII, find the abstract word relation items, analogies, true-false and opposites, easier than the more concrete examples, number series, arithmetic and information. This fact is brought out particularly by the comparatively large percentages of errors in arithmetic and information items. On the other hand, the two "out of school" groups, V and VI, find true-false and analogies to be very difficult items and arithmetic and number series comparatively easy items.

The fact that the information items are the most difficult ones for all the school groups seems to give some verification to the opinions of Mr. Edison and others, that our college graduates do not have a wealth of specific information.

It is our belief that information of value has been gained from the more detailed analysis of mental test records. However, this information must be verified by several independent investigators before it can be used in the service of vocational guidance or industry.

TWO STUDIES OF ATTENTION TO ADVERTISEMENTS

H. K. NIXON

Psychological Laboratory, Columbia University

INTRODUCTION

In this article are reported two studies of attention to advertisements, one having to do with the influence of color and the other dealing with the influence of pictures of people in an advertisement. The data here presented are derived from the results of a research previously published by the author[1] and are intended as an extension of the conclusions of that investigation.

The experiment from which the present results were secured may be briefly described. In general it consisted in presenting two advertisements simultaneously, one advertisement possessing the feature whose attention value was being tested, and the other not possessing this factor, and observing the length of time the subject looked at each advertisement. The full technique of this experimental method and results showing its validity have already been presented[2] and will not be discussed here. Sufficient to say that it serves as an objective approach to problems of attention which have previously been attacked by less direct methods.

The subjects of the experiment were 30 in number, consisting of business men, housewives, and students in Columbia University, 14 men and 16 women. The age range was from eighteen to fifty. In the first experiment here reported thirty-four advertisements having pictures of people and thirty-four having pictures of objects were paired. In the experiment on color forty-four colored and forty-four black and white advertisements were paired. In the first case the factors of position (right or left hand page), color, size of illustration, possession of border, and degree of complexity were kept in balance, so that the pictures of people appeared half of the time on the left hand page, half of the time with color, and so on. In the case of the color experiment similar care was taken. All of the advertisements used were full page, of Saturday Evening Post size. They were bound together in a book of which the subject turned the pages. For the part of the experiment dealt with here the subject spent thirty seconds on each pair, distributing his time between the members of the pair as he was inclined. Unknown to the subject the observer recorded the distribution of eye fixations and the following pages present the tabulation of the data thus derived in the case of pictures of people and of color.

---

[1] H. K. Nixon, Attention and Interest in Advertising, Archives of Psychology, Columbia University, R. S. Woodworth, Editor, 1924.

[2] Loc. cit.

I. THE USE OF PICTURES OF PEOPLE AS AN ATTENTION FACTOR IN ADVERTISING

The practical advertising man has always recognized the value of illustrations of people as a method of attracting attention. A survey of current magazines shows that over 65 per cent of full page advertisements contain pictures of humans as a prominent feature, and it is probable that this use is on the increase. Academic psychology has likewise noted that people are prepotent factors in attention. As far as the writer knows, no study has ever been made to determine with exactitude the real efficiency of illustrations of people. Hollingworth[3] has, indeed, cited an experiment showing that persons and faces are more easily remembered than objects and he has stressed the value of suggested activity, which is likely to be the activity of humans. Aside from this, the topic seems to have been neglected. As the data here presented will indicate, this neglect is serious, for in any experiment on advertisements one might almost as well neglect size or position and might better neglect color or border or size of display type than to neglect this factor. It is, therefore, both from the point of view of the advertiser and that of the experimenter on advertising material, of importance that definite data be secured on this subject. The figures presented in this article are suggestive of what may develop.

---

[3] Hollingworth, H. L., Advertising and Selling, p. 114.

The data of the original experiment of which this work is an elaboration indicated that subjects, in a surprising large number of cases, tend to first turn the eyes to advertisements having pictures of people in them, as opposed to advertisements having only illustrations of objects. For convenience the original figures are reproduced in table 1.

The original data also showed that pictures of people maintained a decided interest throughout the period of exposure.

TABLE 1

First fixations out of possible 34 (Thirty subjects)

	AVERAGE NUMBER	$S.D._{Av.}$
34 advertisements with pictures of people...	18.93	0.48
34 advertisements with pictures of objects...	15.07	0.48

Difference = 3.86. $S.D._{Dif}$... = 0.66.

Averages were worked out showing the accumulated attention value at the end of ten seconds and at the end of thirty seconds. As a general summary these averages were valuable but it seemed desirable to have a more detailed picture of the course of attention. For this purpose the data are so tabulated as to show how much time pictures of people received during each two-and-a-half-second interval throughout the thirty seconds. This is shown in table 2 and is represented graphically in figure 1. Since in the experiment the half-second was used as the unit of time recording all of the figures are in half-seconds.

A consideration of table 2 shows clearly the efficiency of pictures of people as attention factors. In the columns numbered 1 to 12 are given the average times spent by the 30 subjects on the 34 advertisements having pictures of people, calculated for each five half-seconds during the total thirty second exposure. By the method of presentation the group having pictures of objects and the group having the pictures of people had an equal chance to attract attention and had there been no difference in their attention value the averages in each five half-second interval would be 2.50 for each. It will be noted, however, that in most instances the values in table 2 are considerably greater than this, and the averages of the averages range from 2.77 to 3.14. This means that at no time during the thirty seconds were pictures of objects as efficient as their rivals. The small S.D.'s of these averages indicate that they are reliable and by these figures there is not one chance in ten thousand that in any interval the true average would fall as low as 2.50, the point at which the two kinds of illustrations would be equal. It may be said that these results are substantiated by two other experiments which will not be discussed here, in which other groups of subjects and other material gave averages even more favorable to pictures of people.

In figure 1 these facts are presented in graphic form. It will be noted that pictures of people begin with a superior record for first fixations (FF), the curve rising for about twenty-five half-seconds. This indicates that not only does this device attract immediate attention but it serves to hold interest and in this respect is probably superior to any other yet tested. This first period of rise of the curve is followed by a gradual fall, the curve never approaching, however, in the total thirty seconds very much nearer the mid-line than it was at first. If we adopt the somewhat misleading graphic expedient of allowing the area below the curve to represent relative attention to advertisements having pictures of people, the superiority of this group is evident to the eye.

From the original experiment there are also available data showing the memory value of advertisements having pictures of people, as derived from a recall memory test. That value is indicated on the figure by the column at the right. From this it is apparent that the superior attention getting power of these advertisements is reflected in greater frequency of recall.

TABLE 2

Average time spent on 34 advertisements with pictures of people during each 2½ second interval for 30 seconds

(Figures are in half-seconds)

SUBJECT NUMBER	AVERAGE TIME OUT OF POSSIBLE 5 HALF-SECONDS											
	Interval 1	Interval 2	Interval 3	Interval 4	Interval 5	Interval 6	Interval 7	Interval 8	Interval 9	Interval 10	Interval 11	Interval 12
1	2.38	3.00	3.11	2.94	3.32	3.14	2.79	3.32	3.52	3.11	3.02	2.64
2	2.91	3.17	3.02	3.14	3.44	3.14	2.94	3.20	3.20	3.02	3.29	3.32
3	2.79	3.47	3.50	3.41	3.61	3.82	3.61	4.17	4.05	4.20	4.20	4.00
4	2.67	3.05	4.11	3.73	3.47	3.05	2.91	2.55	2.58	2.23	1.73	2.23
5	2.85	2.76	2.58	2.85	2.70	2.52	2.47	2.52	2.41	2.23	2.70	2.58
6	3.23	3.67	3.47	2.94	2.79	2.76	2.41	2.32	2.29	2.26	2.38	2.55
7	2.85	3.67	3.76	3.67	3.52	3.61	3.61	3.38	3.44	3.05	3.29	3.35
8	2.61	2.67	2.79	3.29	3.26	3.05	3.14	2.58	2.85	2.97	3.08	3.02
9	3.20	3.02	3.41	2.94	3.02	3.00	2.55	2.38	1.88	1.82	1.91	2.05
10	2.23	3.11	3.38	3.38	3.23	3.26	3.35	3.17	2.79	2.50	2.73	2.88
11	2.67	2.61	2.52	2.85	2.38	2.73	2.88	2.97	2.64	2.64	2.55	2.64
12	3.11	4.05	4.05	4.02	4.00	3.29	3.20	2.70	2.67	2.29	2.11	2.11
13	3.11	3.35	2.94	2.79	1.97	2.02	2.17	1.94	2.05	2.26	2.97	2.94
14	2.67	2.44	2.23	2.20	2.52	2.58	2.52	2.50	2.26	2.14	2.11	2.14
15	2.85	3.50	3.11	3.38	3.52	3.58	3.14	2.73	2.47	3.23	2.88	2.76
16	2.58	2.61	2.58	2.67	2.26	2.05	2.14	2.38	2.44	2.58	2.70	3.02
17	3.14	3.35	3.88	3.76	4.00	4.00	3.44	3.41	3.23	3.55	3.44	3.20
18	3.44	2.88	2.94	3.26	3.47	3.70	3.44	3.26	3.23	3.23	3.32	3.08
19	2.70	3.44	2.94	2.97	3.32	3.58	3.29	3.26	3.61	3.58	3.52	3.26

20	2.50	2.44	2.35	2.70	2.73	2.94	2.82	2.85	3.11	3.32	3.26	3.05
21	2.97	2.85	2.61	3.41	3.47	3.50	3.50	3.52	3.58	3.52	3.55	3.64
22	3.00	2.32	1.79	2.74	2.70	2.97	3.55	3.35	2.88	2.52	2.58	2.58
23	3.08	2.44	2.91	2.76	3.26	3.47	2.94	2.91	2.55	2.32	2.47	2.14
24	2.47	2.52	2.94	3.20	3.05	2.85	2.70	2.61	2.85	2.88	2.82	2.94
25	2.38	2.64	2.38	2.67	2.91	2.67	2.91	2.91	2.76	2.58	2.64	2.23
26	4.17	4.20	4.32	4.08	4.05	3.79	3.41	2.79	2.79	2.79	2.73	2.82
27	2.50	2.73	2.88	3.23	3.61	2.94	2.44	2.38	2.50	2.67	2.61	2.64
28	2.76	2.70	2.88	2.55	2.82	2.88	3.08	3.38	3.26	2.73	2.73	2.70
29	3.00	3.08	2.61	2.29	2.32	2.76	2.64	2.26	2.38	2.32	2.64	2.32
30	3.67	3.29	3.23	3.32	3.55	3.41	3.38	3.47	3.29	2.79	2.47	2.38
Average ..	2.88	3.03	3.07	3.10	3.14	3.10	2.98	2.90	2.85	2.78	2.81	2.77
$S.D._{Ave}$	0.073	0.081	0.105	0.084	0.097	0.087	0.077	0.088	0.092	0.095	0.094	0.085

It would be interesting to know the real reason or reasons for the advantages shown by these advertisements. We may speculate that it arises from an interest, either innate or acquired as a result of many experiences, in human beings and in all representations of their persons and acts. This is, of course, the

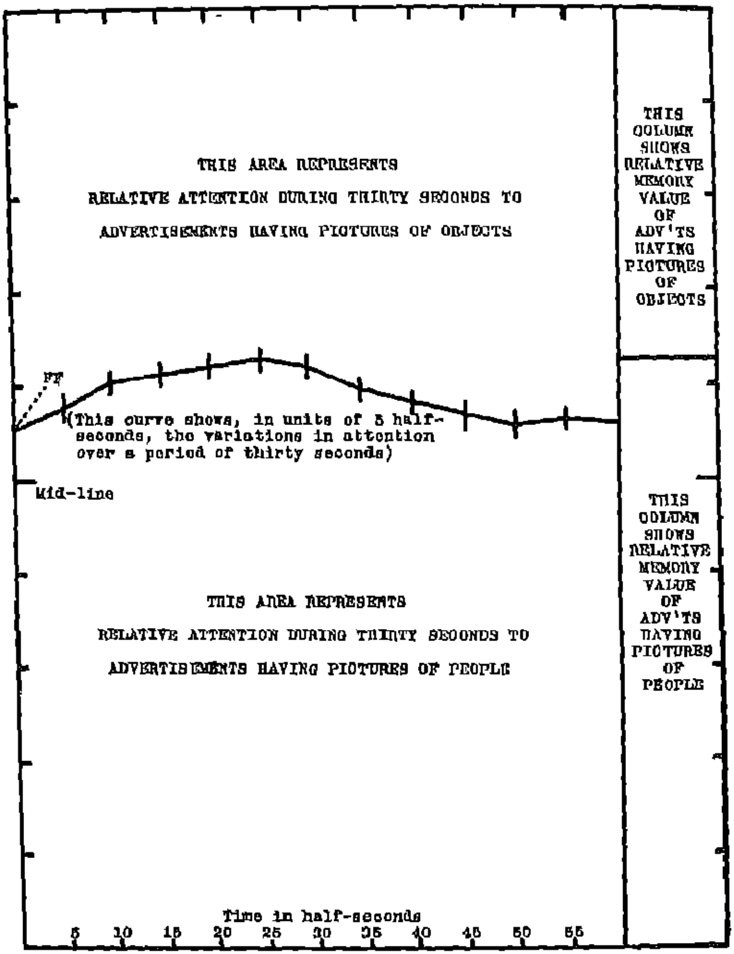


FIG. 1

old theory of "human interest." It seems to the author, however, that there is likely to be more involved than that. For one thing, the situations in which people are presented are *not* so likely to be apparent at a glance. The time taken "to see what it is all about" is likely to be longer and the advertisement *thus gains extra credit. It is likely that many other* factors come in to influence the final showing. It would be foolish, then, to think of a picture of a person as being comparable in simplicity as an advertising device to color or border or type. Much more experimental work is necessary before we can hope to tell exactly why it is that such a picture is more attractive than one of some common object.

From the practical point of view it should therefore be understood that there is no magic by which an indiscriminate sticking of such pictures of people into an advertisement is going to give that advertisement superior attention power. It is *not as simple as putting in color. By a proper use of pictures* of humans the advertiser can very probably add considerably to the attractiveness of any advertisement, but what that proper use is remains for further investigation.

II. THE ATTENTION VALUE OF COLOR IN ADVERTISEMENTS

Franken and Hotchkiss[4] found that colored advertisements had a 13 per *cent greater recognition memory value than an* ordinary black and white page. Kitson[5] reports that the use of color is greatly on the increase since 1912. These *two investigations* are practically the only ones available on this important practical question of the value of the use of color as an attention device. The impression is widespread that color is very effective and millions of dollars are spent every year under the idea that if the advertisement is only given color it is sure to dominate its black and white competitors.

---

[4] *Hotchkiss, G. B., and Franken, R. B., The Attention Value of* Advertisements in a Leading Periodical, published by Graduate School of Business Administration, N. Y. University, 1920.

[5] Kitson, H. D., Minor Studies in the Psychology of Advertising, Journal of Applied Psychology, 6, 1922, pp. 59-68.

TABLE 3

Average time spent on 44 advertisements with color during each $2\frac{1}{2}$ second interval for 30 seconds

(Figures are in half-seconds)

SUBJECT NUMBER	AVERAGE TIME OUT OF POSSIBLE 5 HALF-SECONDS											
	Interval 1	Interval 2	Interval 3	Interval 4	Interval 5	Interval 6	Interval 7	Interval 8	Interval 9	Interval 10	Interval 11	Interval 12
1	3.00	1.90	1.95	2.18	2.02	1.68	1.61	2.02	2.09	2.31	2.38	2.18
2	2.86	2.68	2.65	2.47	2.24	1.90	2.02	2.09	2.06	2.11	2.06	2.09
3	2.58	2.68	2.65	2.54	2.47	2.86	2.65	2.63	2.68	3.02	3.22	3.34
4	2.27	2.31	2.52	2.81	2.72	2.13	2.00	2.00	2.20	1.86	1.93	2.15
5	2.38	1.93	1.86	2.29	1.88	2.15	2.36	2.34	2.45	2.38	2.13	1.97
6	2.65	2.59	2.34	2.24	1.90	1.97	1.93	1.63	1.59	1.81	1.86	2.09
7	2.09	1.70	2.02	1.97	1.93	1.97	1.97	1.93	1.88	2.22	2.45	2.43
8	2.59	2.63	2.59	2.49	2.68	2.59	2.11	2.45	2.22	2.04	2.29	2.49
9	2.49	2.72	3.00	2.29	2.09	1.93	2.11	2.54	2.29	2.09	2.09	2.11
10	2.45	2.47	2.29	2.06	2.29	3.04	2.61	2.31	2.38	2.18	2.20	2.13
11	2.97	2.72	2.84	2.84	2.72	2.74	2.81	2.77	2.84	2.79	2.65	2.81
12	2.36	2.45	2.18	1.95	1.97	1.65	1.63	1.68	1.95	2.59	2.43	2.56
13	1.81	2.04	2.06	2.56	2.24	2.38	2.56	2.49	2.40	2.45	2.77	2.86
14	3.00	2.63	2.40	3.04	3.02	2.88	2.65	2.63	2.88	2.90	2.93	2.97
15	2.93	2.36	2.13	2.18	2.27	2.31	2.20	2.24	2.11	2.36	2.34	2.22
16	2.88	3.09	3.13	2.77	2.15	1.86	1.77	1.77	1.97	1.97	2.00	2.09
17	2.56	2.97	3.09	2.88	3.02	3.06	2.97	2.97	2.95	3.43	3.06	2.79
18	2.79	2.54	2.45	2.52	2.77	2.81	2.81	2.63	2.77	2.45	2.59	2.65
19	3.22	2.63	2.36	2.40	2.38	2.38	2.11	2.00	1.97	2.20	2.13	2.61

20	2.49	2.49	2.49	2.47	2.54	2.29	1.81	2.05	2.15	2.27	2.15	2.11
21	2.38	2.61	2.38	2.36	2.38	2.36	2.20	2.38	2.56	2.72	2.84	3.11
22	2.63	2.65	2.72	2.24	2.54	2.29	2.72	2.70	2.95	2.52	2.61	2.54
23	2.49	2.59	2.54	2.48	2.63	2.65	2.40	2.22	2.20	2.20	2.22	2.31
24	2.72	2.70	2.45	2.24	2.11	2.34	1.68	2.00	1.97	1.95	1.79	1.65
25	3.09	3.02	2.72	2.15	2.22	2.09	2.56	2.72	2.68	3.02	2.84	2.77
26	2.77	2.84	2.74	2.45	2.70	2.54	2.38	2.24	2.34	2.38	2.52	2.59
27	2.56	2.34	2.77	2.68	2.81	2.77	2.93	2.74	2.49	2.36	2.13	2.29
28	2.54	2.54	2.63	2.27	2.77	2.65	2.68	2.34	2.11	2.18	2.45	2.49
29	2.59	2.79	2.93	2.84	2.54	2.36	2.77	2.68	2.68	2.40	2.45	2.27
30	2.56	2.95	2.93	2.72	3.24	3.20	2.81	2.81	2.52	2.86	2.59	2.86
Average ..	2.62	2.55	2.54	2.45	2.44	2.39	2.33	2.33	2.34	2.40	2.40	2.45
$S.D._{Ave}$	0.053	0.058	0.054	0.050	0.064	0.079	0.074	0.064	0.064	0.068	0.063	0.069

The data of this present experiment may serve to throw some light on this question. The results and tabulations were derived

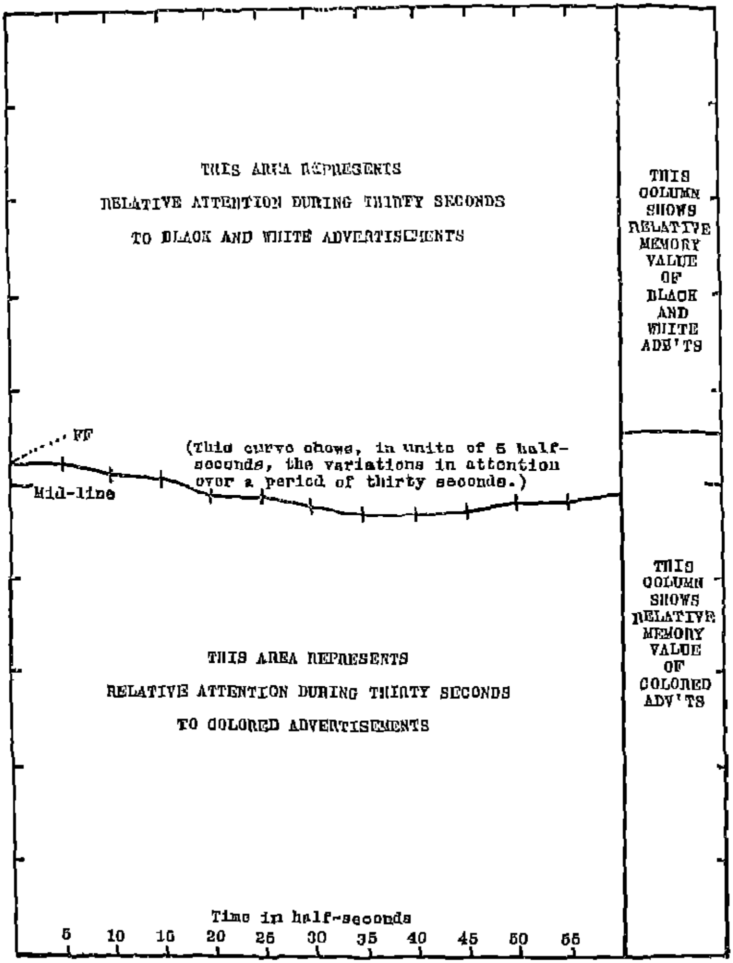


FIG. 2

in exactly the same way as in the case of pictures of people and with the same subjects. They are presented in table 3 and figure 2.

From a consideration of the table and the figure it is apparent that color is surprisingly inferior in attention power. Even in attracting initial attention (FF in the figure), for which it is ordinarily used, it is not nearly as effective as pictures of people. What initial advantage it does possess it loses very rapidly, so that after ten seconds the black and white competitor comes to be slightly superior, as is indicated by the fall of the curve for color below the mid-line on figure 2. Color does, indeed, show considerable memory value, about that found by Franken and Hotchkiss, but in general the use of color as an attention device under modern advertising conditions would hardly seem to be warranted when one considers the great extra cost. The truth is, perhaps, that not color as such but difference or novelty or change are the basic attention factors and in a magazine where colored advertisements prevail any one advertisement need not expect to profit especially in the way of extra attention. This does not invalidate its use for artistic purposes, of course, and from this point of view color may well be worth all it costs. It is likewise probable that in media where the use of color is infrequent its utility may be greater, as may also be the case where it is employed in some unusual or startling manner. In general we would merely urge here that it should not be too blindly accepted as a potent attention device.

CONCLUSIONS

1. Pictures of people are effective attention devices in advertisements. Interest in them increases for a considerable time and they have a high memory value.

2. The results of this experiment show color not to be so effective as an attention device as is generally supposed. It neither attracts great initial attention nor serves to hold interest. Whatever its general utility, in the absence of further experimental evidence it should not be regarded as an especially valuable attention factor under modern advertising conditions.

NOTES ON RATING

PERCIVAL M. SYMONDS

Teachers College, Columbia University

RATING VERSUS RANKING

More and more use is being made of rating scales in personnel work in industry and in education for the quantitative determination of otherwise unmeasurable traits and qualities. Although rating scales are generally preferred, ranking also is sometimes used and has generally been considered the more reliable. By rating I mean the assigning of a position on a scale of already defined steps. Rating is the method used in assigning a value to a composition on the Hillegas Scale or to a sample of handwriting on the Thorndike Scale. Rating is the method we use in the assertion that a child is in excellent, good, fair or poor health, the adjectives used roughly defining the steps on a four-point scale. By ranking I mean the placing of individuals in order as in arranging a gymnasium class in order of height or in sorting the cards in a pack in order from ace to king. If we think of individuals as differing in qualities in certain ways as from little to much, good to bad, desirable to undesirable, we may rank persons in such qualities as leadership, cleanliness, sportsmanship, and the like. At one end we place the individual with little of the trait in mind and at the other end the individual with much of the trait, trying to make each individual in the series exhibit more of the trait than his predecessor.

It is commonly thought that ranking is more reliable than rating because it demands a closer comparison. Allport (Social Psychology, 1924, p. 128) says "The ranking method, available only in rating an actual group all of whom are known to each rater, is *more accurate*. . . . . Comparisons are more concrete and definite than in the scoring method." In rating, comparison is made with roughly defined steps or classes in the scale and the individual is placed in the nearest fitting class according to the opinion of the rater. But in ranking, one individual is compared directly with another.

The following experimentation was made to determine the facts. Two seventh grade teachers rated on a graphic scale a class of forty pupils on seven traits and seven habits. The school was organized on the departmental system so that the class was seen by the two teachers equally long and equally often. After a week the same teachers ranked the same pupil on the same traits and habits. An exactly similar program was carried out by two other teachers in the same school except that the ranking was done first and the rating second to allow for practice effect in our conclusions if there should be any. The traits and habits used were:

Honesty	Copies another's work
Obedience	Conforms to rules
Courtesy	Avoids passing in front of others
Orderliness	Keeps desk in good order
Cleanliness,	Has clean hands
Sportsmanship	Is a good loser in a contest
Promptness	Puts away work quickly

The traits and habits were chosen primarily because of their prominence in school life and their availability for observation rather than for their importance.

The test of the relative reliability of the ratings as against the rankings was to compare the coefficients of correlation of the two teachers working on the same classes on any trait. That is, teacher A correlated with teacher B in *rating* honesty 0.472. The same two teachers correlated in *ranking* the same group in honesty 0.396. The average of twenty-eight coefficients by rating was 0.438 and the average of twenty-eight coefficients by ranking 0.445 (using Pearson's correction to Spearman's ρ which is obtained in using the Spearman rank formula for correlation). The two methods give nearly identical results. In another similar set of ratings and rankings found in another school the same result held—rating and ranking were equally reliable. Also in a college class experiment in rating and ranking the difficulty of words in spelling, ranking had an average correlation with the true difficulties on the Ayres' spelling scale of 0.8966, rating had an average correlation with the true difficulties of 0.8973, and a second ranking a correlation of 0.8842. We may conclude then that rating and ranking are equally reliable.

Knight (1924) says,

> Judges who are presumably equally competent rate the same person as almost wholly honest, half honest, rather dishonest, or an out-and-out cheat. This disagreement is the product of legitimate variations in knowledge about the person and of varying standards of honesty held by the judges themselves. The way to avoid this pitfall is not to use a rating involving concepts of virtues, but a system of ranking persons according to their relative merits. The question should be not, "How honest is he?" but "Is he more or less honest than this person or that?"'

Knight's objection to rating scales is not wholly an objection to their differences in reliability—he is objecting to the tendency to rank high or low on a rating scale as a whole. One teacher might give a class an average rating of 4 on a rating scale in honesty and another teacher give the same class a rating of 2 and yet the two might correlate perfectly. In using the ranking method we can transmute the ranks into sigma indices by means of the tables such as are found in Chu (1922) or into percentiles by means of the tables such as are found in Haught (1923), and can assume that the means of different groups are the same, an assumption which probably is nearer the truth in most cases than taking the means to be as different as may often be found on rating scales. On the other hand, every one that I have asked tells me that rating is more pleasant than ranking. The graphic rating scale seems to be most easy to use. Ranking when the groups become large grows confusing and the effort to make fine distinctions between individuals becomes irksome. Our results show that the fine distinctions simply are not made.

I should like therefore to suggest an *improvement in rating* schemes which would keep the advantages of both methods. Most rating schemes are like this:

RATING CHART A

Pupil's name.................................Date....................
School...Grade...................
Rated by...

Directions for using the Rating Chart

1. *Let these ratings represent your own judgment.* Do not confer with anyone in making them.
2. In each trait or characteristic named below compare this pupil with the average pupil of the same age.
3. In any particular trait disregard every other trait except that one. Many ratings are rendered valueless because the rater allows himself to be influenced by a generally favorable or unfavorable impression which he has formed of the person rated. Do not rate a pupil high on all traits simply because he is exceptional in some. Children are often very high in some traits and low in others.
4. Place a cross somewhere on the line running from "very high" to "very low" to indicate this child's standing in each quality. *You may place your cross at any point on the line.* It is not necessary to locate it at any of the division points or above any descriptive phrase.
5. Do not study too long over any one trait. Give for each the best judgment you can and go on to the next.
6. Give a rating for every trait.
7. The ratings will be held strictly confidential.

HEALTH—Is he generally healthy and vigorous?

Bad	Poor	Average	Good	Excellent

LEADERSHIP—Does he take the lead in school affairs or does he follow others?

Always follows others	Rather tends to follow	Average	Rather tends to be a leader	Masterly, not easily influenced

The usual method with such a scale is to go through the traits for one individual and then turn the sheet to the next individual. Even though instructed to rate all individuals on one trait before turning to the next trait, no direct comparisons can be made since only one page is visible at a time. To make use of the close comparison which is demanded in ranking the following rating scheme is recommended:

RATING CHART B

Rated by...Date.........................
School...Grade.........................

Directions for using the Rating Chart

1. Let these ratings represent your own judgment. Do not confer with anyone in making them.
2. In each trait or characteristic named below compare this pupil with the average pupil of the same age.
3. In this particular trait disregard every other trait except this one. Many ratings are rendered valueless because the rater allows himself to be influenced by a generally favorable or unfavorable impression which he has formed of the person rated. Do not rate a pupil high on this trait simply because he is exceptional in some. Children are often very high in some traits and low in others.
4. Place a cross in one of the compartments running from "very high" to "very low" to indicate each child's standing in this quality.
5. Do not study too long over any one child. Give for each child the best judgment you can and go on to the next.
6. Give a rating for each child.
7. The ratings will be held strictly confidential.
8. Try to let the percentages guide you as to the number of crosses to fill in each compartment.

Trait: HEALTH

Is he generally healthy or vigorous?

PUPIL	4% VERY BAD	11% BAD	21% POOR	28% AVERAGE	21% GOOD	11% VERY GOOD	4% EXCELLENT
Charles..........	....	....	....	....	....	....	....
William..........	....	....	....	....	....	....	....
George..........	....	....	....	....	....	....	....

Such a rating chart keeps the desirable features of the graphic rating chart and also permits the close comparison of ranking.

To be sure the ratings for any one individual are spread over several sheets, but as in any group test the ratings for the several tests for any individual may easily be brought together when the numerical values of the graphic ratings are measured.

Whatever such a chart would contribute toward increased rating accuracy, it would probably in no way do away with the necessity of having three or more qualified and competent persons rate independently in order to have a final pooled rating whose reliability could be depended upon for individual diagnosis.

THE SIZE OF THE HALO EFFECT IN RATINGS

For some time users of rating devices have noticed that ratings of human traits are inaccurate because the rater's judgment of the trait in a certain individual is influenced by his general impression of that individual. So far as I am able to determine Webb in his excellent and only too little read monograph "Character and Intelligence" was the first to notice this halo effect in 1915. On page 70, he says "Let us suppose, for instance, that the observers, in estimating the intelligence qualities, are biased in the direction of marking subjects who possess other desirable qualities too highly, and vice versa." Thorndike (1920) seems to be the first to notice this phenomenon in this country and to use the term halo in identifying it. At that time he said, "The magnitude of the constant error of the halo, as we have called it, also seems surprisingly large, though we lack objective criteria by which to determine its exact size." In this paper partial correlations are used to determine the size of this halo effect.

As previously described, correlations were found between the ratings of the two teachers on any trait or habit. That is, teacher A correlated with teacher B in the ratings on honesty 0.472. Composite ratings of the traits were found by adding together the ratings of the seven traits for each child. This composite was taken as the general impression of the child by the rater. Then partial correlation coefficients of the second order were found for the ratings on each trait with the composites for each of the raters made constant.

1. The halo effect which heretofore has been merely assumed is here demonstrated to be a reality. Seven trait ratings have their correlations raised by 0.245 by the halo effect. Seven habit ratings have their correlations raised by 0.177 by the halo effect.

TABLE I

Coefficients of correlation between ratings of teacher A and teacher B

	ZERO ORDER COEFFICIENTS	COEFFICIENTS WITH COMPOSITES OF A AND B MADE CONSTANT	DIFFERENCES
Honesty	0.471	0.193	−0.278
Obedience	0.387	−0.036	−0.423
Courtesy	0.406	0.110	−0.296
Orderliness	0.190	0.104	−0.086
Cleanliness	0.470	0.553	0.083
Sportmanship	0.360	0.000	−0.360
Promptness	0.448	0.094	−0.354
Average	0.390	0.145	−0.245
Copies another's work	0.587	0.470	−0.117
Conforms to rules	0.313	0.052	−0.269
Avoids passing in front of others	0.306	0.116	−0.190
Keeps desk in good order	0.331	0.110	−0.221
Has clean hands	0.338	0.277	−0.061
Is a good loser in a contest	0.488	0.264	−0.224
Puts away work quickly	0.256	0.101	−0.155
Average	0.377	0.200	−0.177

2. With our data we are not able to demonstrate with statistical sureness that there is a real difference in the rating of traits and habits. But there do seem to be differences which are real. Obedience loses 0.423 with halo made constant while cleanliness gains 0.083 with halo partialled out. The difference is not one of habits versus traits. Possible reasons for a large halo effect in any trait or habit are (1) the trait or habit is one which is not easily observed, (2) the trait or habit is one which is not commonly observed or thought about, such

as one which is not usually emphasized in a classroom, (3) the trait or habit is not clearly defined, (4) the trait or habit is one which involves reactions with other people rather than mere personal behavior, (5) the trait or habit is one with high moral importance in its usual connotation.

The ratings mentioned in this paper were made by four able teachers of Kaahumanu School, Honolulu, Hawaii: Miss Grenville Hatch, Mrs. E. Lemon, Miss Ilda Astleford and Miss E. Astleford.

REFERENCES

ALLPORT, F. H. 1924 Social Psychology.

CHU, J. P. 1922 Chinese students in America: Qualities associated with their success. Teachers College Contributions to Education, no. 127.

FREYD, M. 1923 The graphic rating scale. Jour. Educational Psychology, vol. 14-2, February, 1923, pp. 83–102.

HAUGHT, F. B. 1923 A scheme for combining incomplete rankings. Jour. Applied Psychology, VII-2, June, 1923, p. 172.

KNIGHT, F. B. 1924 Analysis of teaching and teachers. Jour. Educational Research, vol. X-5, October, 1924, p. 224.

PATERSON, D. G. 1923 Methods of rating human qualities in psychology and business. Annals American Academy of Political and Social Science, vol. CX-199, November, 1923, pp. 81–93.

THORNDIKE, E. L. 1920 A constant error in psychological rating. Jour. Applied Psychology, 4: pp. 25–29.

WEBB, E. 1915 Character and intelligence. British Jour. Psychology, no. 3, 1915, pp. 1–99.

A TEST FOR CLASSIFICATION OF STUDENTS IN CHEMISTRY

WINIFRED F. HYDE
AND
J. P. GUILFORD
University of Nebraska

The test here discussed gave such gratifying results that the writers feel justified in reporting their experience with it. The test is the outgrowth of an attempt to solve a problem which was brought to the department of psychology by an instructor in chemistry.

At the above-mentioned university, there are two beginning courses in chemistry: Chemistry 1—five hours credit—a course for those who have had no high school chemistry; and Chemistry 3—three hours credit—for those who have had one year of high school chemistry; both courses to prepare the students for the same advanced work. Chemistry 3 has always had an undue number of failures, many of whom for some reason or other needed more thorough drilling in the fundamentals of the subject, and had they entered Chemistry 1, would have stood a better chance of making a passing mark. According to custom, at mid-semester those who showed need of more elementary work in the subject or a prospect of failure were advised to go into Chemistry 1, but with the inconvenience, in some cases, of changing their semester schedule. The instructor wanted to know if a test could be devised which would detect these students so that they might be advised to enter Chemistry 1 at the beginning of the semester.

Two main factors are probably responsible for failure in Chemistry 3, namely, lack of ability for chemistry, and lack of preparation in high school. Our test, accordingly, contained two parts, Part I, which we will call a Psychological Test, and

Part II, a Chemistry Information Test. The instructor had found through experience that the students are most likely to fail in the solution of chemical problems. These problems are of a rather uniform type, involving the use of ratio and proportion, and requiring the ability to see analogous relationships. Our psychological examination therefore emphasized that type of problem and was composed of four tests:

Test 1, consisting of twenty arithmetic problems with the insertion of a few simple chemical problems which were typical of the kind the student would have to solve in the course.

Test 2, a word analogies test, a very few of them involving chemical terms.

Test 3, a form analogies test, such as is used in the Thorndike test for college freshmen. This presents the same kind of problem as in Test 2, except in the form of non-verbal material.

Test 4, a number series completion test, such as is used in the Army Alpha and other group tests.

The Chemistry Information test was composed of the following:

Test 5, a sentence completion test, the blanks to be filled with chemical terms found in a miscellaneous list at the top of the page.

Test 6, a true-false test, containing twenty statements of well-known facts in chemistry.

The total possible score in the psychological test is 70, weighting the arithmetic test by doubling the number right, and for the chemistry information test, 55 points, or a total combined score of 125. Since the aim was to pick out the least capable chemistry students, the tests were made very simple and the time long enough so that only the slowest could not finish. The total time for work on the tests amounted to about twenty-two minutes.

Two alternative forms were provided, and although an attempt was made to make them of equal difficulty, Form I proved to be a little more difficult in the information part, as will be seen in the table of medians (table 1).

The range was sufficiently great to allow for good discrimination of ability as shown in the distribution curves in figure 1.

The curve for the psychological test is skewed to the right, which makes for better discrimination among those in the lower quartile,—a condition we tried to produce by making the test easy. But the curve for the chemistry information test is skewed in the opposite direction, which makes it more easy to pick out the very superior rather than the very inferior in this test. The test is too difficult for the purpose for which it was intended.

TABLE 1

Medians

	PSYCHO-LOGICAL	CHEMISTRY	COMBINED	NUMBER OF CASES
Form I........................	41.62	14.83	56.33	107
Form II........................	41.55	18.8	60.55	106

But the proof of the value of a test is the pragmatic one: does it work? The correlation of the tests with the semester grades is very significant. It is disturbed by only one fact, that seventeen of the students at the beginning of the semester went into Chemistry 1 upon the advice of the instructor. The latter was given the names of the lowest ten per cent in the two tests, and it was upon this basis and a conference between student and instructor that they were advised to take the more elementary course. But neither the instructor nor any of those who had anything to do with grading the students in Chemistry 3 knew any of the scores of the students until the end of the semester.

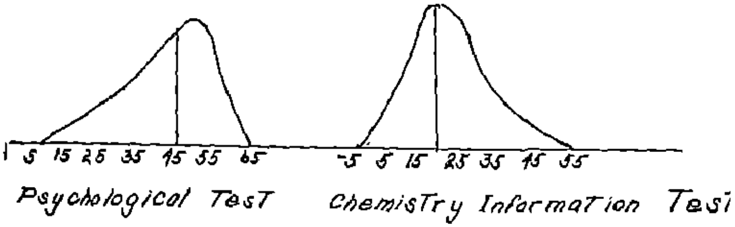


FIG. 1

We have included in our correlations eight of the students who went into Chemistry 1 and failed, assuming that they would have failed also in the more difficult course, but did not include the *nine others who passed in Chemistry 1* because we are not sure that they would have passed in Chemistry 3. The coefficients of correlation are given in table 2.

The correlations would tend to show that the psychological test is a much better criterion of ability to succeed in general chemistry than is the test for chemical information retained from a high school course. This agrees well with the opinion of the instructor, that failure in chemistry is not so much due to lack of preparation in high school as it is to lack of native ability. It also accounts for the fact that eight of the seventeen who made low scores in the test failed although they went into the easier course.

TABLE 2

TEST		CORRELATION WITH GRADE	P.E.
Psychological	Form I	+0.448	±0.0537
	Form II	+0.649	±0.0388
	Both	+0.566	±0.0455
Chemistry Information	Form I	+0.204	±0.0648
	Form II	*+0.437*	*±0.0545*
	Both	+0.331	±0.0590
Combined score	Form I	+0.453	±0.0535
	Form II	+0.583	±0.0435
	Both	+0.517	±0.0493

But coefficients of correlation do not always tell us just how well we can select the potential failure, or what mistakes we might make by selecting an arbitrary passing limit for a test. Since the psychological test correlates better with success in chemistry than does the chemistry information test, and even better than the two combined, it would be wise to take the score in that test alone as a criterion of success in the course. Table 3 shows what per cent of the students were failed or conditioned in Chemistry 3 for each degree of success in the psychological test.

Thus, those who fall below 20 in the test are almost sure to fail, and also half of those between 20 and 30. It would be well, then, to advise all those below 30 to take the more elementary course unless they showed sufficient reasons on other grounds why they are able to pass successfully the more difficult course.

TABLE 3

SCORE	CASES	FAILURE	CONDITION	PER CENT F AND C
				per cent
10-19	3	2	1	100
20-29	18	7	2	50
30-39	48	6	4	21
40-49	76	3	8	14
50-59	37	0	1	3

TABLE 4

STUDENT	SCORE			INSTRUCTOR'S STATEMENT
	Psychology	Chemistry	Total	
C N. T.	47	19	66	Not overly bright
C J. B.	55	26	81	Lazy
C J. M.	44	20	64	Does not work
C J. M.	47	18	65	Lazy. Poor record. Pals with boys who do not work
C R. M.	49	14	63	Really should not have been conditioned. Has worked it off with good examination. Football. Doing well the second semester
C L. M.	44	22	66	No statement
C R. G.	42	24	60	Had been out of school. Hard to adjust himself
C P. B.	47	13	60	"Sluffer." Bright
C A. S.	42	19	61	Outside work handicaps
F R. G.	41	25	65	All right but sickly

We cannot overlook the fact, however, that this method would eliminate only a third of the failures, as can be seen from the table. Of the 22 who made a score above 30 and were failed or conditioned, many reasons other than lack of ability conspired to prevent them from coming up to our expectations. Those who fall between 20 and 30 may be classed as of doubtful ability. But the twelve who made a score above 40 (the median is 42) are surely not lacking in ability. We asked the instructor for statements with regard to these twelve students, and invariably his opinion agreed with our supposition. The statements are given in table 4.

TABLE 5

STUDENT	SCORE			INSTRUCTOR'S STATEMENT
	Psychology	Chemistry	Total	
J. D.	16	4	20	Hardworking, slow
C. D.	24	−5	19	Older student; slow; hard worker
J. K.	18	14	32	Very slow
B. E.	26	11	37	Seems capable; much interested in chemistry

Four students with scores below 30 made a grade of 80 in Chemistry 1. A statement from their instructor (table 5) helps to clear up these cases.

Three of these, because they were slow, were evidently handicapped in the tests, which although allowed a generous amount of time, hurried them too much to be able to do their best work.

Taking all these things into consideration, the tests proved their value to classify students in the two chemistry courses. The question is raised whether, since it is native ability and not preparation in high school which is most essential for success in chemistry, some of the students of very high ability who have had no high school chemistry could not carry the more difficult and briefer course and thus save time in their university work.

Such a plan has not yet been tried, but our results would seem to justify such an experiment. The results also verify the fact that purely psychological or intelligence tests are on the whole a better criterion of success in college than are entrance examinations upon material learned in high school.

CONCLUSIONS

1. By means of a psychological test given at the beginning of the school year and taking not over twenty minutes of time, we are able to eliminate two-thirds of the failures in Chemistry 3, who lack the ability, and to save about half of these from failing in beginning chemistry by placing them in a more elementary course.

2. A psychological test given to freshmen is of more value for predicting success in college than are examinations on high school subjects.

NOTES AND NEWS

FELLOWSHIPS IN THE PSYCHOLOGY AND EDUCATION OF THE PRE-SCHOOL CHILD

The Merrill-Palmer School of Detroit, Michigan, is offering for the year 1925–1926, three fellowships for graduate study in the field of the psychology and education of the pre-school child.

A preference will be given to students who are candidates for higher degrees. Some universities, by special arrangement, are willing to accept work done at this school as part of the requirements for a higher degree. It is entirely possible to do experimental work which will serve as the basis of a thesis for a degree. The fellowships pay $1000 each.

The Merrill-Palmer School offers the following facilities:

Two nursery school centers: One for children between eighteen months and three and one-half years of age, with an enrollment of twenty-five, and the other for children between three and five years, with an enrollment of thirty-five.

A psychological laboratory in which all of the children of both schools are carefully measured and re-measured. This laboratory has now completed the standardization of a new series of *tests for children between eighteen months and five and one-half years*, by half-year periods. The standardization was based upon between five and six hundred children.

A consultation center to which physicians, parents, and social workers refer young children for diagnosis and advice. The center is equipped to make mental measurements, to give complete physical examinations, if these are not already provided for, and to make social investigations. Problems of behavior are analyzed, and advice given about treatment. Parents are advised about the educational environment of young children.

The school is equipped to offer excellent courses in the health and nutrition of young childhood, which may be taken as a minor by students specializing in psychology.

Applications giving full particulars of the academic background and of previous professional experience, if the candidate has had experience in schools, clinics, or social agencies, should be mailed to the Merrill-Palmer School before the end of the present academic year.

HELEN T. WOOLLEY, PH.D.,
Psychologist and Director of
the Nursery School.

Dr. Donald A. Laird, Colgate University, will be the chief psychological editor of *Industrial Psychology*, a new scientific publication which will make its appearance in January, 1926. Mr. Percy S. Brown, works manager of the Corona Typewriter Company and President of the Taylor Society will be chief industrial editor. Each issue will contain 80 pages of reading matter devoted to the problems of employment, personnel, industrial relations and other problems connected with industry. Hopeful but impossible notions are to be culled by having both editors pass on each paper before publication. A special rate of $4.00 a year is offered to those subscribing before the first issue. Address *Industrial Psychology*, Box 34, Hamilton, N. Y.

The JOURNAL OF APPLIED PSYCHOLOGY extends a cordial welcome to this new contribution.

The Eighth International Congress of Psychology will be held at Groningen, Holland, from September 6 to 11, 1926. Invitations have been sent to about 200 well-known psychologists. Others who desire to attend are requested to apply to the First Secretary, F. Roels, 86 Maliebaan, Utrecht, Holland.

On May 14 to 16, 1925, at Ann Arbor, was held a meeting of the Taylor Society, an international society to promote the science and the art of administration and of management, under the auspices of the University of Michigan. Of special interest was a paper presented by C. S. Yoakum, Professor of Personnel Management, University of Michigan, on Experimental Psychology in Personnel Problems. This was followd by discussion led by Prof. A. W. Kornhauser, School of Commerce and Administration, University of Chicago, and Lillian M. Gilbreth, Frank B. Gilbreth, Inc., Monclair, N. J.

Dr. Harry D. Kitson has accepted a position in the Department of Psychology at Columbia University. He has been connected for some years with the Department of Psychology, University of Indiana, Bloomington, Ind.

Dr. Ada Hart Arlitt, formerly of Bryn Mawr College, now with the Cincinnati Central Clinic, will become head of a Department of Child Care and Training to be established next year in the University of Cincinnati.

BOOK REVIEWS

G. M. Ruch. *Improvement of the Written Examination.* Scott, Foresman and Co. x + 193 pp. Price $1.80.

During the past two decades there has been a growing amount of experimental data on the reliability—or rather the unreliability—of teachers' marks. The general conclusion common to all the studies reported in the technical literature is unanimous—examination grades and semester grades vary widely not only from one instructor to another, *but the grading of any instructor is highly unreliable.* In other words, the measure of relative school success or failure most commonly used is in point of fact so largely a chance affair that the proper placement of *a given individual pupil is bound to be grossly in error* in a large proportion of the cases. The measurement of the achievement of educational goals is by the same token subject to the same limitations insofar as these measurements are a function of the typical examination grade.

All of this is a matter of a b c's for the educational technician. Yet the fact remains that very little has been offered by way of remedy. Wood, Monroe, McCall, Toops, Knight are a few of the names that come to mind relative to those who have given some constructive suggestions. What has been needed is a book which would summarize the results of the necessarily piecemeal activities of the investigator, evaluate them critically, analyze the aims of education relative to examinations, and *map out a program about which the garden variety of teacher could do* something. Standardized tests, useful as they proved and will continue to be, do not answer the need. Apart from other weighty considerations, they are not adapted to the relatively short time intervals at which successive measurements are desirable.

The book here under review performs a notable service in line with the above considerations. Dr. Ruch has not only admirably summarized heretofore scattered conclusions, but adds a considerable amount of experimental data on his own account to show the relative merits of different types of examinations. The theoretical and technical discussion is well within the grasp of the average classroom teacher and is marked by scientific conservatism rather than any tendency to enthusiastic overstatement. There is no attempt to buttress absence of experimental data with dogmatism. The educational practioner in the classroom will be most grateful for the inclusion of directions for preparing adequate types of examinations together with much illustrative material in various kinds of subject matter such as chemistry, biology, literature and composition, French, Latin, physics, spelling, reading, history, and manual training.

The six chapter headings will indicate the content and organization of the book: Chapter I, Functions of Written Examinations; Chapter II, The Criteria of a Good Examination; Chapter III, Sources of Error in Written Examinations; Chapter IV, Types and Construction of the Newer Objective Examinations; Chapter V, Experimental Studies of Several Types of Objective Examinations; Chapter VI, Statistical Considerations Related to Examination Technique. The Appendix contains a complete high school content examination as well as some of the illustrative material referred to above. The summaries and questions at the close of each Chapter are exceptionally well worked out.

In the judgment of the present reviewer the book should find a place in all elementary courses in educational tests and measurements as well as in teachers' reading circles, study groups, etc. To quote from the Preface by Dr. Ernest Horn: "There has now developed a new type of test program. . . . . These new and more perfect methods of testing fill a need which cannot be met by the use of standardized tests. . . . . Dr. Ruch has performed a real service in presenting the evidence concerning the reliability of various forms of tests and in illustrating so clearly the different ways in which these tests may be prepared. There is no question . . . . but that this book marks out the next big field of improvement in testing the results of teaching."

H. H. REMMERS,
Purdue University.

JOHN MONK SAUNDERS AND GEORGE PALMER. *Brain Tests.* G. P. Putnam's Sons: New York, 1925. Pp. 95. $1.50.

Adieu to crosswords! Here we find a small book, with pencil attached, by which we can test our brains and ascertain just how our mind works. Part of the jacket resembles the crossword cover but the book itself is filled with abbreviated psychological tests which could not have been made more intriguing.

A key and an "average" score is given for each test in the back of the book whereby the reader may check up his powers of observation, range of information, ability to follow directions and tread through a maze, make change or try his skill at geometric figures.

I predict considerable success for this book, and think that psychology will receive a little more popularization from it. What it will tell the reader about himself is difficult to say, but he will certainly find it interesting.

DONALD A. LAIRD, PH.D.,
Colgate Psychological Laboratory.

NATALIE KNEELAND. *Cases in Retail Salesmanship.* Merchandise Manuals, A. W. Shaw Company, 1924. Pp. 189. Price $1.50.

Any one who can imagine the ignorance and trepidation with which a young salesgirl approaches her first customer in a department store can understand how great is the need for aids for retails salespersons. A number of aids are already available. There are several textbooks on retail selling besides the long list of books covering salesmanship in general. Courses are offered by correspondence, and training is offered by most department stores. In addition to these general helps for salespersons there are a number of books giving technical information about the goods sold in department stores. In spite of the volume of such aids there is still room for another kind of help. Salespersons find it hard to pass general principles of selling to concrete instances. And they find the masses of technical information about goods quite beyond their power to assimilate. Recognizing these difficulties, Doctor Charters and his associates in the Research Bureau for Retail Training, of Pittsburgh have prepared a series of sales manuals that have as their aim the presentation of facts in their concrete setting and the recommendations of practises in the concrete.

The number of manuals announced by the publisher to be in preparation is seventeen in all. The first one, "Cases in Retail Salesmanship," is an attempt to apply the case method of instruction to all the departments of a department store. It is a collection of 135 accounts of actual sales, some reported by students of the Bureau, some by shoppers. They are reported in dialog form. The sales were selected as illustrative of certain types of situations such as "Knowing what is in stock;" "Selecting merchandise appropriate to the customer's taste;" "Knowing when to approach the customer;" "Handling two customers at once." For example,under the heading, "Finding out what the customer wants by asking questions," appears the following dialog:

Salesperson—Is there something for you, madam?
Customer—Well, I was just trying to decide.
Salesperson—Are you looking for yourself?
Customer—No, I was thinking about a dress for my daughter.
Salesperson—Does she require a girl's or junior size?
Customer—She is just my size.
Salesperson—We have some lovely new sport silk frocks. Would you care for something like this?
Customer—Oh, that's just what she wants. I heard her speak of having one, the other day.
Salesperson—Is she dark or light?
Customer—Her hair is red.
Salesperson—Oh, then this same dress in tan and brown would be lovely on her. Isn't that a lovely color? (showing dress).

Customer—And how pretty she would look in it! She loves tan and brown so much.

Salesperson—The girls are wearing these dresses so much now for sport wear.

Customer—I believe I'll let you send that up.

At the close of each account questions are presented which are intended to lead the student to analyze the sale and determine its strong and weak points. References are also given to several books on retail selling.

Several errors crept into this manual: On page 30 "fallacy" is misspelled; on page 33, yoke; on page 52 the question, "When will this be out?" should be attributed to the customer instead of to the salesperson; on page 178 appears the phrase "principles . . . . was;" on page 183, "iridescent" is misspelled.

NATALIE KNEELAND. *Waists.* Merchandise Manuals. A. W. Shaw Company, 1924. Pp. 126. Price $1.50. *Hosiery, Knit Underwear and Gloves.* Pp. 126. Price $1.50.

FREDONIA JANE RINGO. *Girls' and Juniors' Ready-to-Wear.* Merchandise Manuals. A. W. Shaw Company, 1924. Pp. 160. Price $1.50.

The manuals devoted to various commodities are compilations of the facts about the commodities which a salesperson would find most useful in "talking up" her goods. For pedagogical reasons they are arranged in the form of questions and answers, such as:

Question: "What different qualities of cotton are used in underwear?"

Answer: "There are many grades of qualities of cotton. The most important to know are the Sea Island, Egyptian, American Upland, and Peruvian."

Then follow paragraphs on each of these. Each section is ended by "Test Questions." A section in each manual is devoted to Methods of Teaching, much of the material in this section is the same in all the manuals.

A unique and much-needed section of the manual on Girls' and Juniors' Ready-to-wear contains a vocabulary for describing values. A "black-list" shows words that salespersons should avoid, such as cheap, classy, clever, cute, elegant, grand, sweet, etc. Instructions are given to use words conveying definite ideas instead of general ideas. For example, instead of saying, "This is a lovely blue," one should say, "This is Delft blue, or cadet blue." A long list of words that may be used in describing texture includes airy, clinging, filmy, fleecy, sheer, supple, etc. Words describing different styles are artistic, conservative, draped, fitted, flaring, jaunty, ornate, severe, swagger, ultra, etc. For the sake of this section alone every retail salesperson should study these books. The practical value of all the aids offered by this series of manuals will undoubtedly insure their wide acceptance by those who are in charge of training retail salespersons.

HARRY D. KITSON,
Indiana University.

Helmholtz's Treatise on Physiological Optics. Translated from the Third German Edition. Edited by James P. C. Southall. Vol. I. The Optical Society of America, 1924. Pp. xxi, 482.

At a meeting of the Optical Society of America in 1921 it was proposed to commemorate the hundredth anniversary of the birth of Helmholtz "in a more useful and substantial way by bringing out the long-delayed English translation of the Physiological Optics, and accordingly a special committee was appointed to have charge of this business." Let it be said at once that the Optical Society, and more especially Professor Southall and his collaborators in the arduous task of translation and editing, have performed a notable and valuable service by publishing in English "one of the four or five greatest monuments of human genius in the scientific line."

The French have had their translations of Helmholtz sooner than we have. The *Handbuch der physiologischen Optik* first made its appearance, in four installments, between the years 1856 and 1866. Within a year it was published in French by E. Javal and N. Th. Klein. *Die Lehre von den Tonempfindungen* appeared in 1862 and was translated by G. Guéroult six years later and supplemented by an appendix after the third German edition in 1874. But in biding out time we have come into possession of much more serviceable versions of Helmholtz's works. The present edition of the *Optik* by Professor Southall, even if for no other reason than that it is a translation of the third German edition, is more valuable than the French one. And A. J. Ellis' version of the *Tonempfindungen* is in a class by itself.

Professor Southall and his committee decided that the third German edition was beyond question the one to be reproduced in English. This edition was brought out in three volumes after Helmholtz's death, under the editorship of W. Nagel, in collaboration with Professors A. Gullstrand and J. von Kries, between 1909 and 1911. Nagel followed the text of the first edition instead of the second, which meant the exclusion of certain of König's researches but gave, on the other hand, a more faithful picture of the extraordinary brilliance with which Helmholtz dealt with phenomena which took him into the fields of physics, mathematics, physiology, psychology, and philosophy. The edition was expanded to nearly double the size of the original by the addition of new material in the field of physiological optics contributed by the editors.

Readers of the English translation, then, will have all of this material before them, and in addition, significant new contributions which have been made to the subject in the last decade and numerous excellent foot-notes by Professor Southall and his collaborators.

The first volume, Anatomical Descriptions and Dioptrics of the Eye, contains, in addition to all the material of the German edition, an entirely new chapter by Professor Gullstrand on his own remarkable researches in ophthalmoscopy; a bibliography of more recent literature on anatomy of the eye compiled by D. H. Hooker which, "although it does not pretend to be a complete list by any means," has nearly 200 titles; numerous new foot-notes and references to literature; and a portrait of Helmholtz which appeared in the third volume of the German edition. In the preparation of the first volume the editor was assisted in large measure by D. H. Hooker and L. D. Weld.

The English-reading scientific world is under a debt of gratitude to the American Optical Society for the appearance and very excellent preparation of Helmholtz's great work, and especially to Mr. Adolph Lomb without whose "continual advice and encouragement the achievement would never have been accomplished."

Helmholtz's Treatise on Physiological Optics. Translated from the Third German Edition. Edited by James P. C. Southall. Vol. II. Published by the Optical Society of America, 1924. Pp. x, 480.

The second volume of the third German edition of the *Optik* appeared in 1911, one year after the third volume. The delay was due to the illness of the editor, W. Nagel. After Nagel's death von Kries assumed the responsibility for the completion of the second volume. Nagel had already prepared parts of an appendix for the second volume on adaptation, twilight vision, and the duplicity theory, which were included by the new editor together with an appendix of his own on normal and anomalous color systems and theories of vision. These additions by Nagel and von Kries to the original text of Helmholtz are reproduced, of course, in the English translation of the second volume which now follows close upon the appearance of the first volume.

Psychologists will find the contents of the second volume, sensations of vision, of more immediate interest to them than those of the first volume. Here one has Helmholtz's great researches on visual stimulation, simple and compound colors and their attributes, contrast, etc., and the appendices by Nagel and von Kries. The English edition contains, in addition, a chapter by Dr. Christine Ladd-Franklin on the nature of color sensations. In this chapter Dr. Ladd-Franklin has discussed certain of the weak features of the Helmholtz theory of vision, reproduced the facts and diagrams derived from König's important researches, which are not elsewhere fully treated in the English translation since the editors decided to follow the third rather than the second German edition, and given an account of her own development theory of color sensations. The English volume also contains three entirely new notes especially prepared by von Kries in simultaneous contrast, the connections between various forms of color vision in dichromats, and methods for preparing the luminosity of light of different colors; and an impressive bibliography, compiled by the editor, of works relating to sensations of vision which have appeared since 1911.

What has already been said about the preparation of the first volume is equally true of the second. Professor Southall and his collaborators deserve the highest praise for bringing before English readers such an excellent version of Helmholtz's monumental work! In the preparation of the second volume the editor was assisted by H. Laurens, M. Dresbach, L. T. Troland, and E. J. Wall.

C. C. PRATT,
Harvard University.

NEW BOOKS AND PAMPHLETS RECEIVED[1]

Books and pamphlets for review should be sent to James P. Porter, Department of Psychology, Ohio University, Athens, Ohio.

Abnormal Psychology and Education. FRANK WATTS. D. Appleton and Company, 1924.

Challenge of Childhood, The. IRA S. WILE. Thomas Seltzer, 1925.

Coöperation in Adult Education. ELLEN C. LOMBARD. Home Education Circular No. 6. Bureau of Education, Washington, D. C. 34 pp.

Genetic Studies of Genius, Vol. I. LEWIS M. TERMAN. Stanford University Press, Stanford University, Calif. Price $5.00. 848 pp.

Graphic Methods in Education. J. HAROLD WILLIAMS. Houghton Mifflin Company, 1924. Price $2.00.

Human Conduct and the Law. MARY C. LOVE. George Banta Publishing Company, Menasha, Wisconsin. 318 pp.

Lefthandedness, A New Interpretation. LUIS PARSON BEAUFORT. The Macmillan Company, 1924.

Legal Provisions for Rural High Schools. WILLIAM R. HOOD. Bulletin No. 40, 1924, Bureau of Education, Washington, D. C. 60 pp.

Problems in Physical Education. JAMES FREDERICK ROGERS. Physical Education Series No. 5. Bureau of Education, Washington, D. C. 19 pp.

Psychotechnik der Beroepskeuze. J. L. PRAK. Bij. J. B. Wolters U. M., Groningen, Den Haag. 272 pp.

Purposive Speaking. ROBERT WEST. The Macmillan Company, 1924.

References on Education for Citizenship. Library Leaflet No. 30, January, 1925. Bureau of Education, Washington, D. C. 16 pp.

Statistical Method in Educational Measurement. ARTHUR S. OTIS. World Book Company, Yonkers, N. Y. 337 pp.

Statistics of State School Systems. FRANK M. PHILLIPS. Bulletin No. 31, 1924. Bureau of Education, Washington, D. C. 42 pp.

A Study of 260 School Consolidations. J. F. ABEL. Bulletin No. 32, 1924, Bureau of Education, Washington, D. C. 39 pp.

Ways of the Mind, The. Charles Scribner's Sons, 1925. 336 pp.

Yiddish Press, The. An Americanizing Agency. MORDECAI SOLTES. Teachers College, Columbia University. 1925.

---

[1] Mention here does not preclude further comment.

Studies from Columbia University, Teachers College

Changing Conceptions Relative to the Planning of Lessons. LOIS COFFEY MOSSMAN. No. 147, 1924.

Content and Form in Tests of Intelligence. EDWIN MAURICE BAILOR. 1924.

A Comparative Study of the Mental Capacity of Children of Foreign Parentage. MAY BERE. No. 154, 1924.

Detailed Factors in Latin Prognosis. ORLIE M. CLEM. No. 144, 1924.

Diagnostic Study of the Subject Matter of High School Chemistry, A. SAMUEL RALPH POWERS. No. 149, 1924.

Evaluation of English Literature in the High School. CHARLES SUMNER CROW. No. 141, 1924.

Measurement of Educational Need, The. PAUL R. MORT. No. 150, 1924.

Some Values Derived from Extensive Reading of General Science. FRANCIS DAY CURTIS. No. 163, 1924.

Study of Intelligence Test Elements, A. LEONA VINCENT. No. 152, 1924.

Tentative Inventory of the Habits of Children from Two to Four Years of Age, A. RUTH ANDRUS. No. 160, 1924.

Values of New Type Examinations in the High School. STERLING G. BRINKLEY. No. 161, 1924.

THE RELATIVE TRANSFER EFFECTS OF SUPERVISED PLAY AND FORMAL GYMNASTICS[1]

CONSTANCE E. DOWD AND ADA HART ARLITT

Psychological Laboratory of the Vocation Bureau and Central Clinic, Cincinnati, Ohio

This investigation, begun in June, 1923, and completed in August of the same year, was undertaken in an attempt to determine the relative effects of a daily period of training in formal gymnastics versus the same amount of time spent in supervised play.

While there have been exhaustive studies of the extent and direction of the transfer of training in many fields, the only study which deals with the same problem as our own, namely transfer in the field of formal gymnastics, is an investigation made by Williams.[2] Williams' subjects were children from two fourth grades in the Horace Mann School.

Two groups were selected that were equal in mental ability, that is, the median intelligence quotient for one group was 114 while that for the other was 113. The range in intelligence quotients for the one was from 96 to 148 while that for the other was from 100 to 146. These groups were then equalized for motor ability, posture and growth and development as tested by the tests listed below. An obedience test of Williams was also given to the group The final equalized groups had 14 in each group for tests of skill, 18 in each group for the obedience test and from 16 to 19 in each group for the growth and development tests. The tests are listed by Williams as follows: height, weight, and lung capacity, measured in the usual manner, girth of chest normal and expanded, strength of right forearm and strength of left forearm, Creek (a test of William's own which is too long to describe but which involves mainly balancing), Bean Bag (throw), Thrusting (aiming) Stilts, Balance Beam, and two so-called tests of Alertness which involved crossing out dots and crosses.[3] After the groups had been equalized the experimental group was given two periods weekly of formal gymnastics for a period of eight months. The control group spent the same amount of time weekly in plays and games. As the same instructor worked with both groups the groups did not receive the training at the same hour of the same day. At the end of the training period the initial tests were repeated. The means of the gains for the play and game group were larger than those for the formal gymnastics group in all tests but three, chest expansion, strength of left forearm and thrusting. The difference on the whole favored the play and game group but there was practically no reliability of the difference for any tests but alertness. Williams used McCall's Expermental Coefficient $\left(\frac{\text{difference}}{2.78 \times \text{S. D. difference}}\right)$ as a way of expressing reliability. The average Experimental Coefficient for all of the tests listed except posture and obedience is 0.30. That is to say the difference in favor of the play and game group is 0.30 as large as it shoud be in order to be practically certain that the methods used for the play and game group were better than those used for the formal gymnastics group. There seemed to be no marked transfer to posture. In obedience as measured by Williams' test the play group improved more than did the formal gymnastics group. The Foster heart test seems also to have been used but the results from this test were not treated fully.

---

[1] Read before Section Q of the American Association for the Advancement of Science, January 1, 1925.

[2] Williams, J. F. A Comparative Study of Formal Gymnastics and Play for Fourth Grade Children. Teacher's College Record, Vol. XXIII, September, 1922, pp. 327–360.

Our experiment differs from that of Williams in that: (1) The conditions under which our tests were given were somewhat more controlled; (2) the training periods for our groups occurred at the same hour of the day; (3) the tests used in our experiment differed from those used by Williams in many instances; (4) the hygiene of rest, sleep and nutrition which Williams speaks of as being difficult to control were as completely controlled as was possible for a normal group. The routine activities of the day, the hours of sleep and rest, the amount and kinds of food given to the subjects and the kind and amount of exercise taken were kept equal as nearly as possible for all subjects. (5) Our subjects received intensive training for five weeks while in Williams' experiment the training was given in regular school periods, we presume twice a week, for the school year. The subjects used in the present study were thirty-five girls living in a summer camp in Maine. The subjects ranged in age from 7 to 18. The median chronological age for both groups was 13.22.

---

[3] The parentheses are the writer's.

The following procedure was used to divide the total group into two equal groups. The Pintner Non-Language test was given to the group as a whole. All subjects received this test on the same afternoon. Nineteen tests of motor ability were also given to the whole group but these were administered individually. The tests of motor ability were chosen with a view to selecting two groups which would be equal at the outset of the training period in the measurable activities on which formal gymnastics is said to have a noticeable effect, namely, speed, accuracy or precision of movement, kinaesthetic discrimination, balance, strength, and general bodily coordination.[4] As tests of speed we used the Woodworth and Wells cancellation test, card sorting, rate of right and left hand tapping (with the ordinary tapping board), and a 75-yard dash. For Kinaesthetic discrimination we used twenty 2-inch cubes weighted with shot to vary from each other by 0.5 gram. The lightest weight weighed 20 grams the heaviest weighed 30 grams. The constant was the middle weight. The subject was blindfolded and the constant and the variable were presented. The order was varied to prevent any tendency on the part of the subject to select the weight in accordance with the order in which it was presented. The difference between the constant and the variable was increased until the subject could choose the heavier of the two correctly 8 out of 10 times. The same procedure was followed with the constant and the series of weights decreasing in heaviness until the subject could choose the lighter weight correctly 8 out of 10 times. The procedure was repeated in each direction going from the lightest weight to one which could just be discriminated as lighter 8 out of 10 times and from the heaviest weight to one which could just be discriminated as heavier 8 out of 10 times. For balance we used a balance line similar to a walking boom and hopping in circles with increasing distances between the circles. The circles were one foot in diameter and varied in distance from each other from 3 inches to 2 feet 6 inches. The subject's score on the balance line was computed on the basis of the distance traversed without loss of balance, her score on the circles was the actual number of circles into which she could hop without loss of balance or touching the line which marked the outside edge of the circle. There were 10 circles in all and the distances between the circles were increased by three inches each starting from the first circle which was just 3 inches from circle 2. Circle 3 was 6 inches from circle 2, circle 4 was 9 inches from circle 3 and so on. For bean bag throw, the first test of accuracy or precision of movement, we used 5 concentric circles. The outside circle was 5 feet in diameter, the next was 4 feet, the next 3, the next 2 feet and the inside circle 1 foot in diameter. The subject stood 10 feet from the line marking the outside circle and threw 5 bean bags. All bags thrown into the innermost circle gave the subject a score of 10, those in the next circle scored 8, in the next 6, the next 4 and the outside circle 2. The subject's score was the total score made with the five bags. Target aiming with a fencing foil was also used as a test of precision. The subject stood at a distance of

---

[4] The tests are listed under these headings largely as a matter of convenience. Though functions other than those under the headings under which they are subsumed are also involved.

4 feet from a target consisting of 5 concentric circles differing from each other by 2 inches in diameter. The outside circle was 10 inches in diameter, the inside circle was 2 inches in diameter. The score for touching within the radius of a circle differed from 10 points for touching the inside circle to 2 points for touching the outside circle. The circles were different colors in order to make it easy for the subject to differentiate between them clearly.

As the strength test we used the hand dynamometer. The score on this was the better of two trials with each hand. As tests of general bodily coördination we used Baseball Throw for distance, Basket Ball Throw for distance, Hop, Step, and Jump, Running High Jump, Standing Broad Jump, and Running Broad Jump. We also determined the vital capacity of each subject by the use of a Wet Spirometer.

The nineteen tests described above were given to the whole group during the first week of camp. They were given to each individual in a prescribed order and the time of day and the conditions under which the tests were given were kept constant for the group, and noted down in order that the series given at the end of the training period should be given under the same conditions as the initial tests.

When all of the test scores were in, a scale was devised for each test which would give scores that could be added and averaged for all of the 19 tests. On the scale for each the highest score was rated 100 and the lowest score was rated zero. The intermediate scores were rated according to their position on the scale. If L equals the lowest raw score and R equals the range of raw scores, then the reduced score for any particular raw score S is represented by the following equation:

$$\text{Reduced score} = (S - L) \times \frac{100}{R}$$

That is, the lower limit of the range was taken as zero and the upper limit as 100 and each individual was given his proportional place in the range. A reduced score of 40 means that the raw score was 40 per cent of the distance from the bottom to the top of the range.[5]

This method was chosen chiefly because of its simplicity and because all of the scoring had to be completed and the groups divided within ten days after the opening of camp. No attempt was made to weight tests as we have no data on which to base weighting of motor tests for children.

Reduced scores of all tests for each subject were averaged and this average was used as the final score for motor ability. The mental test scores were used separately as a second criterion for division. The whole camp was then divided in such a way as to distribute chronological ages, mental test scores and motor test scores as evenly as possible. The method of pairing was used as far as was possible with so small a group. Individuals were shifted about until the groups were practically equal in average motor test score, average mental test score, and average chronological age. The average scores for the group were as follows:

	FORMAL GYMNASTICS GROUP	SUPERVISED PLAY GROUP
Average motor score	55.61	54.52
Average mental test score	61.04	62.41
Average chronological age	13.22	13.23

---

[5] The formula actually used was $100-\left(H - S \times \frac{100}{R}\right)$ where H = highest score. That is we began at 100 at top of range and measured down instead of beginning at the lower end of the range and measuring up. Both formulae are the same since $R = H - L$.

$$
\begin{aligned}
\text{Reduced score} &= 100 - \left(H - S \times \frac{100}{R}\right)\\
&= 100\left(1 - \frac{H - S}{R}\right)\\
&= 100\left(\frac{R - H + S}{R}\right)\\
&= 100 \times \left(\frac{H - L - H + S}{R}\right)\\
&= 100 \times \left(\frac{S - L}{R}\right)\\
&= (S - L) \times \frac{100}{R}
\end{aligned}
$$

Training was begun on the tenth day after the opening of camp. The Formal Group was given a twenty minute period of formal gymnastics every morning at 9:00 for a week and thereafter a thirty-minute period for four more weeks. That is to say, the class was held every morning except Sunday for five weeks. When a subject missed one of the weekday periods this was made up on Sunday morning. The exercises given were the regular Swedish System exercises as described in Enebuske's "Progressive Gymnastic Days Orders"[6] and in Skarastrom's "Gymnastic Teaching."[7] The theory of progression as set forth in Skarstrom was kept in view. In the opinion of the instructor the group showed marked improvement in execution and in ability to perform more difficult exercises by the end of the training period. The class was held out of doors in fair weather and in a large, well ventilated room in rainy weather.

Beginning on the same day and at the same hour the Supervised Play Group had play and games each day for the same number of minutes as the Formal Group had formal gymnastics. The periods were spent in playing games such as Dodge Ball and Relay Races and in performing stunts such as are described in Pearl and Brown in "Health by Stunts."[8] The stunts were performed in the most informal way, Each girl tried most of them but made her own choice as to which of them she would practice. Practice was thus to some extent introduced but it was in such a hit or miss fashion as to be in no way allied to drill of the formal period. As examples of the stunts the following are representative: chinning, push up, head stand, hand stand, hawk dive and balancing with chairs.

At the end of the training period the initial tests were repeated under the same conditions and in the same order as they had been given in the original series.

The results of each test are set forth in table 1. Gains in raw score are averaged for each group and the difference between the average gain for each group in each test is expressed in terms of the probable error of the difference. In each case where the Supervised Play group is superior to the Formal Gymnastics Group the number is preceded by a minus sign.

---

[6] Enebuske, C. J. Progressive Gymnastics Day's Orders. Silver, Burdette & Co.

[7] Skarst. Gymnastic Teaching, Amer. Phys. Ed. Assoc., Springfield.

[8] Health by Stunts. Pearl and Brownin. MacMillan Co. 1921.

TABLE 1

*Difference in average gains of experimental and control groups**

	AVERAGE GAIN		DIFFERENCE	P.E.	DIFFERENCE IN P.E. UNITS
	Formal Gymnastic Group	Supervised Play Group			
Cancellation	7.88	8.05	0.17	1.27	0.14
Card Sorting	4.50	4.79	0.29	1.42	0.20
Tapping, right	9.33	8.06	1.27	3.24	0.39
Tapping, left	1.22	6.16	4.94	3.99	1.24
Weight Lifting	−0.03	0.23	0.26	2.72	0.97
Balance Line	0.05	1.35	1.30	0.77	1.68
Bean Bags	−3.22	−7.29	4.07	1.86	2.18
Hopping	0.61	0.41	0.20	0.72	0.28
Target	2.94	3.05	0.11	1.20	0.09
Vital Capacity	4.11	2.41	1.70	2.21	0.77
Hand Grip, right	2.89	5.47	2.58	1.76	1.46
Hand Grip, left	0.61	2.41	1.80	1.37	1.31
Baseball Throw	−63.88	11.02	74.90	40.46	1.85
Basket Ball Throw	−6.11	16.76	22.87	15.36	1.40
Dash	0.00	0.08	0.08	1.34	0.01
Hop, Step and Jump	17.50	10.26	7.24	4.65	1.55
High Jump	3.04	3.35	0.59	0.77	0.76
Running Broad Jump	*20.66*	*22.94*	*2.28*	*2.49*	*0.91*
Standing Broad Jump	2.61	1.76	0.85	1.04	0.81

* For suggestions as to statistical procedure the writers are indebted to Professor R. S. Woodworth.

Gains are stated in terms of the unit of measure for each test, for example, in time for card sorting, in score for target aiming and in inches for the jumps.

In twelve tests the average gain for the Formal Gymnastics Group was less than that for the Supervised Play Group. These tests are Cancellation, Card Sorting, Tapping left hand, Weight lifting, Balance line, Target aiming, Right Hand Grip, Left Hand Grip, Baseball Throw, Basketball Throw, 75 yard dash, and Running Broad Jump. The Formal Gymnastics Group was superior to The Supervised Play Group in Tapping, right hand, Bean Bag Aim, Hopping in Circles, Vital Capacity, Hop, Step and Jump, and Standing Broad Jump.

It is interesting to note the type of tests in which the Supervised Play Group showed superiority. In Basketball throw and Baseball throw, both of which are held to involve general bodily coördination, the Supervised Play Group was superior but in Bean Bag throw the formal Gymnastics Group made the better score. In weight lifting, a test involving kinaesthetic discrimination, and in card sorting which involves speed of perception and speed and accuracy of movement the Supervised Play Group were better than the Formal Gymnastics Group, though formal gymnastics is supposed to have a definitely beneficial effect in these directions.

Before going into a prolonged discussion of the differential effects of the two types of training the authors were interested to determine whether such differences as were found between the experimental and the control groups were reliable. The P.E. of the difference was computed for each test. A glance at the table will show that in the case of all tests except Bean Bag throw, and Hop, Step, and Jump which show a difference in favor of the Formal group and in Baseball throw, Basketball throw, and rate of Tapping with the left hand which show a difference in favor of the Supervised Play Group, the differences are so small as to be practically negligible.

After the tests had been treated separately all of the differences in P.E. units for all of the tests were averaged. The final score was the probable error of this average. This was 0.28 in favor of the Supervised Play Group. The P.E. of this average was 0.175 which shows that there is an even chance that the difference might be as small as 0.105 and there is some fair chance that the difference might be practically zero, or even that the Formal Gymnastics Group might show a slight superiority over the Supervised Play Group.

Lest it be said that our results are open to question because the subjects used in this experiment were in camp and therefore exercising their muscles, and that formal gymnastics could therefore not be expected to have as much effect as if the children had been in a school where there was little opportunity for other exercise, we should state that the hypothesis is that formal gymnastics is of more value than other forms of exercise. It is well agreed that it is better than no exercise at all. If formal gymnastics have a better and a more specific effect than mere exercise, then one-half hour daily for four weeks following a period of twenty minutes daily for one week should have showed more effect than the same amount of time spent in play and games. All functions, and practically all muscle groups, which have been the subjects of transfer experiments have been in use in some other way while the transfer experiment was going on. Formal gymnastics is supposed to give certain definite effects, as for example increased precision of movement, which mere exercise does not give. These effects should have been obtainable whether the subjects were having other exercise or not.

The writers are not in a position to state to just what extent the improvement in either of the groups was due to the mere fact that the subjects were in a summer camp.

In view of the results of the experiment we should state that under the conditions of this experiment Formal Gymnastics given for a period of 20 minutes a day for a period of one week, followed by a period of one-half hour daily for four weeks produced no greater effect than the same amount of time spent in play and games. If anything the Play Group was slightly superior to the Formal Gymnastics Group but the difference is so small as to be practically negligible.

TENDENCIES IN THE USE OF HEADLINES IN MAGAZINE ADVERTISING

C. J. WARDEN, W. A. RORISON AND S. B. SOLINGER

From the Psychological Laboratory of the University of Wisconsin

This report covers briefly the results of an investigation of headline usage in magazine advertising during the twenty year period beginning with the year 1900. The *Saturday Evening Post* was selected as the representative medium on account of its acknowledged leadership in the advertising field. An analysis of the headline was made for the years 1900, 1905, 1910, 1915, and 1920, the data for each year being compiled from the complete issues of the months of March, June, September, and December, taken as representative of the advertising practice of the year. The study naturally falls into two major divisions, the first dealing with certain mechanical aspects of the headline, and the second with the factors of content and form.

I. MECHANICAL FEATURES

Length of headline, length of words used in the headline, the kind of type used (whether upper or lower case) and the size of type used were the four mechanical features studied, and the data gathered will be presented topically in the above order. Only half-page and full-page advertisements were included in this survey.

1. Length of headline. Taking the period as a whole there seems to be no very marked or consistent tendency to shorten the headline, although the modal length of three words for the year 1920 is perhaps significant. The differences in range, average, or mode are less than one might expect from a casual inspection of advertising sections of the *Post* for the years concerned.

2. Length of words used in headlines. In order to determine whether or not the tendency among advertisers of the *Post* was toward the use of longer or shorter words in display lines, the letter units of each word were counted and the average computed. It is obvious from the accompanying table that shorter words were used, more commonly than longer words, throughout the period. Words of over 9 letters rarely occurred and monosyllable words were usually preferred, except when

TABLE 1

Length of headline (number of words)

	1900	1905	1910	1915	1920
Range	3 to 8	1 to 8	2 to 11	1 to 16	1 to 10
Average	6	4	6	5	4
Mode	4	4	5	4	3

TABLE 2

Length of words used in headlines

	1900	1905	1910	1915	1920
Range (letters)	2 to 16	1 to 12	1 to 14	1 to 16	1 to 12
Average	6	5	4	4	4

TABLE 3

Use of upper and lower case type

	1900	1905	1910	1915	1920
	per cent	*per cent*	*per cent*	*per cent*	*per cent*
All—capital	30	25	10	8	4
All—lower case	20	5	5	2	1
Combination	50	70	85	90	95

TABLE 4

Size of type (in points)

	1900	1905	1910	1915	1920
Range of size	18 to 72	18 to 96	12 to 72	12 to 90	12 to 102
Modal size	30	36	36	32	36

the name of the product or firm was used in the display line. The average length of word employed in the headline decreased from six to four letters during the twenty-year period.

3. The use of upper and lower case in headlines. The classification includes the all-capital, the all-lower case and the combination upper and lower case headline. The latter class includes, for the most part, the capitalization of the initial letter of the first word, or of all the more important words of the headline. It will be noted that the straight upper and lower case headline has now been practically eliminated in the half-page and full-page advertisements of the *Post*, although both kinds were used rather extensively in 1900. The combination type of heading has come to be used almost exclusively.

4. Size of type. The point system was used in measuring the type and the values appearing in the accompanying table are in terms of this unit of type size. Type larger than 72 points is rarely found in the display lines of any magazine advertisement. There has been an increase in both the upper range and average since 1900, but not a very large one. The preference seems to be for the 36 point size, at present, with the 42 point size as second choice.

II. THE USE OF DIFFERENT CLASSES OF HEADLINES

In contrast to the mechanical features of the headline, the content and form are factors that may also be used as the basis of interesting comparisons. Following the classification suggested by Starch (Prin. of Adv., pp. 504–509) a survey was made to determine the relative frequency of the various types. The advertisements were first arranged according to size into four groups as follows: full-page (including double-page spreads), half-page, quarter-page, and smaller (from 1 inch to quarter-page). Each group was then analyzed as to content and form, and the classification reduced to the percentages appearing in table 5. The *Saturday Evening Post* was again used, exactly the same issues and years being selected as in the study of mechanical features, advertisements of all sizes being included in this case, however.

TABLE 5

The use of different classes of headlines

	1900	1905	1910	1915	1920
Full-page including spreads					
	per cent	*per cent*	*per cent*	*per cent*	*per cent*
Name of article	0	6	40	36	34
Name of firm	0	12	1	2	1
Selling point	0	12	24	27	30
Indirect heading	0	41	24	21	24
Blind heading	0	17	2	2	1
Command	0	6	5	8	6
Question	0	6	4	4	4
Half-page					
Name of article	100	70	29	16	40
Name of firm		0	2	1	0
Selling point		12	23	38	24
Indirect heading		12	32	28	24
Blind heading		6	2	2	1
Command		0	8	10	6
Question		0	4	5	5
Quarter-page					
Name of article	92	44	40	33	49
Name of firm	0	8	1	1	2
Selling point	0	8	21	19	24
Indirect heading	8	28	28	35	14
Blind heading		4	2	0	7
Command		4	6	9	1
Question		4	2	3	3
Smaller spaces					
Name of article	40	44	62	44	38
Name of firm	12	11	1	0	0
Selling point	20	13	12	32	6
Indirect heading	16	20	14	12	12
Blind heading	4	3	1	1	0
Command	5	6	5	8	29
Question	3	3	5	3	15

The name of the article advertised seems to have been used as the headline most often during this period, regardless of the size of the advertisement. The percentages show a considerable variation from year to year with no very definite trend.

Evidently the firm name was never very popular as a display caption, its use from 1910 and onward being almost negligible in the case of both large and small advertisements.

A fairly consistent increase appears in the use of an explicit selling point in the headings of advertisements of the quarter page size and larger. In the smaller size no trend is indicated.

The indirect heading stands next to the name of the article as a leader in display lines. On the whole 1920 registered a decrease in this type of heading by *Post* advertisers.

The blind heading has very properly grown less and less popular even in the case of the smaller advertisements.

The command and question forms held consistently throughout the period within the 3 to 9 per cent range for quarter-page advertisements and larger. However, a very large increase in the use of both forms, but especially the command, evidently occurred between 1915 and 1920, for the smaller spaces.

PERSONALITY DIFFERENCES IN INTELLIGENCE TEST PERFORMANCE

GROUP INDIVIDUAL DIFFERENCES: THE SIGNIFICANCE OF "OMISSIONS"

GRACE E. MANSON

University of Michigan

INTRODUCTION

Psychologists, especially those who have concentrated upon standardizing and interpreting intelligence tests, are becoming more and more convinced of the fact that intelligence test scores need to be supplemented by other standardized measures if high correlations are to be obtained between test results and criteria of success. This need becomes especially urgent if the criterion is something other than academic grades. Witness the low correlations listed in many of the experiments where the criteria have been success in commerce or industry.[1]

On all sides there is a great demand for reliable measures of "character" and "personality" traits. Diligent researches are being made of methods of analyzing and measuring some of these seemingly elusive factors. The method of experimentation is in a measure determined by the trait under consideration. Written answers may be adequate for some traits, other motor activities as "tasks" may be necessary for other traits. For some time, at least, a battery of tests measuring the same trait will be more apt to give reliable results than any one test alone.

With the economy of time and subjects as a goal, the writer has been lured by the possibility of isolating and measuring certain personality traits as demonstrated in intelligence test performance. This is neither an original nor startling purpose. Wm. M. Brown,[2,3] with the same aim in view, has attempted to isolate the "caution" factor in subjects taking the Thorndike intelligence examination for high school graduates, The Trabue completion test, the Otis examination, or a Binet test. Dr. Brown's paper is suggestive and valuable. The fact that this "caution" factor increases the predictive value of his regression equation for scholastic ability but little, is not evidence that his technique is poor or that "caution" is not a factor in scholastic standing, but rather that "caution" is but one of the many factors contributing to scholastic success. A truer measure of the reliability of his method of measuring "caution" would be obtained, as Dr. Brown suggests, if his "caution" indices were correlated with other measures of caution. But, unfortunately, other measures were not available.

---

[1] Kornhauser, A. W., and Kingsbury, Forrest A., Psychological Tests in Business, The University of Chicago Press, 1924.

THE PROBLEM

Early last year, the writer, having the same conviction as Dr. Brown, that much more information could be gained from the group intelligence test than the total score indicated, began working with Bureau of Personnel Research Test VI. This is a spiral omnibus adaptation of Army Alpha. There are in all 184 questions, 20 arithmetic, 40 opposites, 40 analogies, 24 true-false, 40 information and 20 number series. Questions are not weighted but those requiring more time, namely, the arithmetic, number series and true-false, are fewer in number: 20, 20 and 24, for arithmetic, number series and true-false, as against 40, 40 and 40 for opposites, analogies and information. Following the army scheme, errors on opposites and true-false are doubly penalized. In scoring, omissions

---

[2] Brown, Wm. M., A Study of the "Caution" Factor and Its Importance in Intelligence Test Performance, The American Journal of Psychology, 1024, Vol. 35, No. 3, pp. 368–386.

[3] Brown, Wm. M., Character Traits as Factors in Intelligence Test Performance, Archives of Psychology, May, 1923, No. 6.

on all tests count as errors but are not doubly penalized in the true-false and opposites test.

Preceding the test proper are two pages of directions in which each type of problem is explained by use of examples. Instructions are uniformly given to "work rapidly." The instructions read, "If you are not sure about something, guess at it and go on to the next thing. There are "no catch questions." You will be allowed fifteen minutes. You probably will not be able to finish in the time allowed, but do as much as possible." Then in a line to itself, "Do not skip about. Items skipped will be counted wrong." Four minutes are allowed for the reading of directions, fifteen minutes for the test proper.

This test is not comparable to the Thorndike intelligence examination for several reasons. First of all, the Thorndike test was made especially for high school graduates, particularly for those desirous of entering college, a selected group, whereas Test VI was designed for adults of varying ages and interests. In the Thorndike test, especially in those parts used by Dr. Brown, the subject is encouraged to omit. "If you do not know, go on to the next," whereas in Test VI the subject is told to guess and not to skip. "If you are not sure about something, guess at it and go on to the next thing." "Do not skip about. Items skipped will be counted wrong."

This strong command to guess and not to skip, which is given in the directions for Test VI, led us to wonder why there were so many omissions. Mere misunderstanding of the directions would not account for all of the omissions. It seemed more likely that the omissions were controlled by some selective processes which were stronger than the urge to obey the instructions. To find these selective factors becomes the problem of this paper.

SUBJECTS

In the endeavor to determine the reasons underlying omissions on Test VI, we have studied eight groups of persons who had taken the test. These groups differed in age, in educational status, and in vocational and occupational interests, as well as in final scores obtained on the intelligence examination. Our purpose, in analyzing the intelligence test records of these eight groups, was to see if the six types of tests, arithmetic problems, opposites, analogies, true-false, information, and number series, were of approximately the same relative difficulty for each of the eight groups. The results of this study, which are given in detail in a previous paper,[4] show that there are group differences in the relative difficulty of the several types of questions.

In the present study, we have limited ourselves to two groups of persons included in the previous investigation. These two groups are alike in that (1) the members of each group were graduate students at the time that the test was given; (2) they made about the same final scores on Test VI; and (3) they found the several types of tests comprising the total examination of about the same relative difficulty. These two groups, designated as Groups I and II,[5] were as follows:

Group I consisted of 14 graduate students at the University of Pittsburgh, registered in the Research Bureau for Retail Training.

Group II contained 25 members of a course in mental measurements at the University of Michigan Summer School. The majority of these students were either college or normal school graduates. The two groups differed in vocational interests in that the members of Group I were preparing to become executives in the personnel departments of department stores, whereas the majority of the members of Group II were interested in teaching or administrative positions in public education.

Table 1 gives the average final scores of the two groups and the average number of the last example attempted, together with the standard deviations of the distribution of scores.

---

[4] Manson, Grace E., Group Differences in Intelligence Tests: The Relative Difficulty of the Types of Questions, The Journal of Applied Psychology, June, 1925, Vol. 9, No. 2, pp. 156–175.

[5] Group II is the same as Group VII in the previous paper.

METHOD

Having obtained two groups which, on the surface, seemed so much alike, especially in respect to the difficulty of the intellectual stimulus, Test VI, we then attempted, by further analysis, to find means of differentiating between the two groups and between members of each group in respect to one personality difference, namely, "caution." The method of doing this was to compare (1) the ratios of the total number of errors to the total number of omissions; (2) the percentages of omissions and the number of persons omitting on each type of question; (3) the percentage of errors and the number of persons making errors on each type of question.

By this technique we found how groups and members of groups differed in making a choice between omitting difficult questions or attempting to answer difficult questions, even if it were evident the answer was likely to be wrong. In other words, we had a measure of the readiness to guess.

TABLE I

Average final scores and average number of last example attempted

	GROUP I	GROUP II
Final scores..............................	124.2 σ21	121.8 σ19
Number of last example attempted.......	140 5 σ19	137.8 σ20

The results of this statistical analysis were compared with the introspective reports of the subjects as to why they made each omission.

A scoring method was then devised which aimed to measure certain motor activities which may be described as cautious behavior. The introspective reports were of great value in pointing out the several motives for this cautious behavior as well as in giving insight into the particular stimulus or stimuli which led to cautious behavior in particular subjects. The comparison between the objective records of speed, errors, and omissions and the introspective reports also gives interesting personality clues.

RELATION OF TOTAL ERRORS TO TOTAL OMISSIONS

Table 2 gives the average number of errors and the average number of omissions made by each group.

The sum of errors and omissions is practically the same for the two groups but Group I has a tendency to omit rather than to make errors. The graduate students preparing for business careers were more ready to discount the directions "not to omit" than were the public school group.

TABLE 2

Average number of errors and omissions

	GROUP I	GROUP II
Omissions	3.8	2.2
Errors	12.5	14.0
Sum of errors and omissions	16.3	16.2

TABLE 3

Average percentage of omissions and number of persons omitting

TYPE OF PROBLEM	GROUP I			GROUP II		
	Per cent of omissions	Number of persons omitting	Per cent omitting	Per cent of omissions	Number of persons omitting	Per cent omitting
Arithmetic	0.09	8	0.57	0.019	5	0.20
Opposites	0.005	2	0.14	0.01	2	0.08
Analogies	0.0025	1	0.07	0.000	0	0.00
True-false	0.000	0	0.00	0.000	0	0.00
Information	0.052	7	0.50	0.04	7	0.28
Number series	0.038	4	0.29	0.24	8	0.32
Number of persons in group		14			25	

RELATION OF OMISSIONS TO THE TYPE OF PROBLEM

Table 3 gives the average percentage of omissions on each type of question and the number of persons omitting. This percentage was obtained by dividing the total number of examples of each type of question covered by each subject, into the total number of questions of that type omitted.

Examination of table 3 shows that (1) members of Group I omitted more often than members of Group II; (2) true-false, analogies and opposites were very seldom omitted by either group; and (3) that information items, arithmetic, and number series were more frequently omitted by both groups.

RELATION OF FORM OF THE ANSWER TO NUMBER OF OMISSIONS

What are the reasons for these differences? Is the form of the answer the controlling factor? Are omissions less prevalent in questions where the guessing possibilities are greatest? For example, in opposites and true-false items, the chances of guessing correctly are even. The choice is between scoring one of two words, as:

cautious—heedless. same opposite
Earth is mined coal the from. True false

In the information items and analogies there is a choice of the correct answer from among four possibilities as:

Calcutta is a city in Egypt China India Japan
feather : float :: rock : ages hill sink break

The number series and arithmetic questions do not offer any choice between answers. There is no visible clue to the correct answer. There is only a blank space in which the subject is directed to write an answer, after the mental operations are completed.

If the form of answer were the controlling factor, we should expect the order of omissions to be (1) opposites and true-false, (2) information and analogies, and (3) arithmetic and number series. This order holds except for information items and analogies. Information items are omitted more often than number series, and analogies are not omitted any more frequently than opposites or true-false items.

RELATION OF ERRORS TO THE TYPE OF PROBLEM

If the form of the answer is not the only reason for omissions, what might be another one? Is the difficulty of the question an important cause? Are there many errors on types of questions where there is a number of omissions? Table 4 gives the percentages of error and the number of persons making errors on each type of question.

Information items are the most difficult items for both groups, as is shown by the high percentage of errors and the large number of persons making errors. Opposites and arithmetic problems are the next hardest for Group I, if percentage of errors is taken as a standard; arithmetic and number series, if the number of persons making errors is used as the standard. Arithmetic and number series are the next most difficult items for Group II.

TABLE 4

Average percentage of errors on each type of question and number of persons making errors

TYPE OF PROBLEM	GROUP I			GROUP II		
	Per cent of errors	Number of persons making errors	Percent of persons making errors	Per cent of errors	Number of persons making errors	Percent of persons making errors
Arithmetic	0.09	11	0.70	0.10	19	0.76
Opposites	0.10	8	0.57	0.10	13	0.52
Analogies	0.04	8	0.57	0.03	15	0.60
True-false	0.068	7	0.50	0.054	11	0.44
Information	0.16	14	1.00	0.16	24	0.96
Number series	0.07	10	0.71	0.13	21	0.84
Number of persons in group		14			25	

RELATION OF ERRORS TO OMISSIONS ON THE SEVERAL TYPES OF TESTS

When we compare errors with omissions, we find that errors and omissions are apt to be concentrated on the same type of problems. For example, in each group errors and omissions are few for analogies and true-false statements, whereas both errors and omissions are more numerous in information items and number series.

But Group I is more apt to omit problems that appear difficult or time-consuming than Group II. This is shown particularly in the case of arithmetic problems.

If we consider the sum of the errors and omissions, as a measure of difficulty, each group has an 18 per cent deduction from the number of arithmetic problems covered. But Group I reacted by omitting half of the tedious problems, whereas Group II only omitted 2 per cent and made errors on 16 per cent. Looked at from the point of view of individual subjects, 8, or 57 per cent of the 14 subjects in Group I, omitted one or more arithmetic problems, whereas only 5, or 20 per cent of the 25 members of Group II, omitted any arithmetic examples. The same tendency in Group I, to omit rather than to make errors, is also seen in the case of information items and number series. Fifty per cent of the members of Group I omitted one or more information items. Twenty-eight per cent of Group II omitted one or more information items. Number series are apparently harder for Group II than for Group I. The sum of errors and omissions is 15.4 per cent for Group II, 10.8 per cent for Group I. Yet, Group I has a slightly higher percentage of omissions than Group II. The percentage of persons omitting is about the same, 30 per cent for Group I, 32 per cent for Group II. These comparisons are shown in table 5.

A further evidence of the relative difficulty of the questions is shown in table 6 which gives the number of persons in each group making neither errors nor omissions on each type of question.

Here again, opposites, analogies, and true-false appear easier than arithmetic, information and number series. Therefore we conclude, in answer to our previous question regarding the relation between errors and omissions, that the difficulty of the question, as well as the form of the question, controls the number of omissions.

SUMMARY OF STATISTICAL FINDINGS

Examination of the statistical results shows that: (1) few omissions are made on true-false items because the form of the answer is favorable to guessing, and the material of which the test is made is not difficult for these groups; (2) few omissions are made on analogies because the material is not difficult, even though the form of the questions is not as favorable to guessing

TABLE 5

Percentages of omissions and errors and number of persons omitting and making errors

TYPE OF PROBLEM	GROUP I				GROUP II			
	Per cent omissions	Number omitting	Per cent errors	Number making errors	Per cent omissions	Number omitting	Per cent errors	Number making errors
Arithmetic	0.00	8	0.00	11	0.019	5	0.16	19
Opposites	0.005	2	0.10	8	0.01	2	0.10	13
Analogies	0.0025	1	0.04	8	0.000	0	0.03	15
True-false	0.000	0	0.068	7	0.000	0	0.054	11
Information	0.052	7	0.16	14	0.04	7	0.16	24
Number series	0.038	4	0.07	10	0.024	8	0.13	21
Number of persons in group		14		14		25		25

TABLE 6

Number of persons making neither errors nor omissions

TYPE OF PROBLEM	GROUP I	GROUP II
Arithmetic	1	4
Opposites	4	11
Analogies	6	10
True-false	7	14
Information	0	1
Number series	3	2
Number in group	14	25

as the true-false items; (3) a slightly higher percentage of omissions is found on opposites because the content is more difficult, even though the form of the answer is favorable to guessing; (4) information items are frequently omitted because the material is not as well known. The form of the answer, although moderately favorable for guessing, does not completely counteract the tendency to omit difficult items; (5) number series are not as frequently omitted as information items, although the form of the answer makes correct guessing improbable; percentages of error on number series are not as high as on information items; and (6) arithmetic questions are more frequently omitted by Group I than any other type of question; arithmetic questions are difficult for Group I but they do not make as high a percentage of error on arithmetic problems as on opposites.

Evidently Group I has been influenced by the form of the question and preferred to omit the problems rather than attempt to work them, since the form of the answer makes correct guessing practically impossible. Group II has not omitted but has worked the arithmetic problems, making a large percentage of errors. It is interesting, here, to notice that the sum of percentages of errors and omissions on arithmetic problems in Group I is 18 per cent, in Group II, 17.9 per cent.

INTROSPECTIVE REASONS FOR OMISSIONS

But why do persons omit when the directions tell them not to omit and urge them to guess? What reasons do the persons omitting give? Will these reasons tally with the above evidence gained from the statistical findings?

In order to see if subjects who omitted did so intentionally and for the same reasons that the statistical results seemed to indicate, the writer asked the members of Groups I and II to tell why they left out the particular problems which they did. The verbatim replies follow.

GROUP I

Reasons for omissions

Subject 1. Omitted three arithmetic problems.

"I first tried the problem by algebra—then arithmetic and thereby became so confused that I knew neither."

"From past experience I knew that in a test of this sort, arithmetic is not weighted when scored and that it takes a greater amount of time than another type of problem giving equal credit."

"I did not guess because in such arithmetical problems one does not guess readily."

Subject 2. Omitted two arithmetic problems, two information items, two opposites and one number series.

(Arithmetic)

"I skipped numbers 127 and 130 to save time."

(Information)

"I skipped numbers 42 and 122 because I was not sure of the correct answer."

(Opposites)

"I skipped 153 because I was not sure of the correct answer, and 121 I overlooked."

(Number series)

"I skipped 156 to save time."

Subject 3. Omitted one information item.

"I skipped 153 because I was unfamiliar with the second word and my eye had caught the next question. My impulse was to guess on it but time was called before the impulse could reach my paper."

Subject 4. Omitted one arithmetic, and six information items.

(Information)

"I think in general I omitted questions on general information. Some of these I had never heard of—and it seemed to me better to omit than to make a wild guess which would probably be far from hitting the point."

(Arithmetic)

"I omitted the arithmetic question because I am so dreadfully slow in doing arithmetic."

Subject 5. Omitted one arithmetic and two information items.

"None omitted. Reason—Attempted to follow specific directions to that effect."

Subject 6. Omitted three arithmetic and two information items.

(Information)

"Skipped numbers 11 and 26 because I positively did not know the answer."

(Arithmetic)

"Skipped numbers 111, 127 and 130. Started No. 111, decided the arithmetic probably wasn't weighted in scoring and as they took much more time than the rest, it would be best to skip them. Worked 145 because it was apparent at a glance."

Subject 7. Omitted none.

"Did not skip any because it was counted as an error. Guessed with the possibility that I might get it right. Interested in the problems and did not desire to skip those. Like to work with figures."

Subject 8. Omitted none. No reasons given.

Subject 9. Omitted none.

"I didn't skip any of the questions. First, the directions definitely said not skip. Second, in problems after taking time to read them, it would seem wasted if I didn't complete the operations. It seems to me the time is balanced there. "Third, on questions of knowledge if I don't know instantly, I guess."

Subject 10. Omitted none.

"When these tests were given at the (another) University we were told not to dare to omit any. I had never scored a test, but now I think I would leave out what I didn't know and go on and make time. Almost all my guesses were wrong on this test."

Subject 11. Omitted three information items.

"I skipped because I was sure I didn't know."

Subject 12. Omitted one arithmetic, two information and one number series.

(Information)

"I had nothing to go on in guessing the answer. Where I guessed in other questions I had some idea to direct me."

(Number series and arithmetic.)

"I worked at it and couldn't get it at first and expected to come back and try the problem again after I had done other questions in order to forget my various attempts."

Subject 13. Omitted five information, one arithmetic and one number series.

"I skipped a couple of the problems because I get confused when I cannot get it immediately and cannot get over it in going onto the next question.Therefore, I would rather omit one and answer more of those that I can. I waste time thinking over those I do not know and probably won't get them right in the end anyway."

Subject 14. Omitted eight arithmetic problems, five information items.

"Knowing that score depended on number correct or amount of ground covered I purposely omitted those involving arithmetic calculation intending to go back if time permitted. I was out for speed and did not know it was the purpose of the test to have us complete each problem before proceeding to the next."

The reasons given show the effects of a certain familiarity with test taking and scoring. However, since all members of this group were in the same course in mental measurement, the information was a common factor. The two outstanding reasons for omitting are: (1) to save time; and (2) to avoid committing one's self by putting down answers which are known to be wrong. The arithmetic problems were usually omitted to save time.

It will be recalled that Group II made a smaller percentage of omissions, particularly on arithmetic items. What reasons did they give for omissions?

GROUP II

Reasons for omissions

Subject 1. Omitted eight information items.
"Not knowing I simply did not bother to guess. I do not like to guess."
Subject 2. Omitted three information items.
"I did not know the ones I omitted."
Subject 3. Omitted one number series.
"Overlooked. I probably got in too big a rush."
Subject 4. Omitted none.
Subject 5. Omitted none.
Subject 6. Omitted none.
Subject 7. Omitted none.
Subject 8. Omitted none.
Subject 9. Omitted twenty. No reason given. (absent)
Subject 10. Omitted one. No reason given. (absent)
Sub ect 11. Omitted one number series.
"I didn't knowingly omit any. I don't know why I omitted number nine."
Subject 12. Omitted none.
Subject 13. Omitted one number series.
"Worked at such a nervous tension under the time limit instructions that I felt I'd get along better if I didn't take time to think."
Subject 14. Omitted one information.
"Evidently overlooked. Haste probably caused skip, careless."
Subject 15. Omitted none.

Subject 16. Omitted one number series.

"I thought it would take too much time to figure it out, so I thought I would rather omit it than have it wrong."

Subject 17. Four omissions, three information, one number series.

"Number five evidently overlooked. Since the others were going to be counted wrong, I preferred to have it known that I knew that I did not know."

Subject 18. Omitted none.

"I didn't disobey instructions."

Subject 19. Omitted two number series, one arithmetic.

"Discovered on Army Alpha that I couldn't solve mathematical series of difficulty, so I quit wasting time on them in Test VI as soon as they reached a certain difficulty."

Subject 20. Omitted none.

Subject 21. Omitted one information, one arithmetic.

"The information I omitted because I overlooked it. In the other one (arithmetic) I thought that I could do better by going ahead than stopping to work this problem. I did not want to put down what I knew to be an incorrect answer.

Subject 22. Omitted three, one number series, two analogies.

"No. 105 (number series) didn't know, No. 108 (analogy) stage fright, No. 110 (analogy) hurried."

Subject 23. Omitted one arithmetic.

"I didn't know the answer at a glance."

Subject 24. Omitted two arithmetic, one number series, one information.

"I was careless in writing answers. Then I was careless in following orders."

Subject 25. Omitted none.

"I obeyed instructions."

In comparison with Group I, the reader will notice that the emphasis is upon following instructions. There is little deliberate skipping in order to gain time. The chief reasons for omissions are "overlooked" or "did not know and therefore did not care to commit myself." Since the majority of the items overlooked were either information or number series, items which were difficult for the group as shown by the errors, it seems likely that some of the "overlooking" was not due to chance alone.

We have given in tables 7 and 8 the detailed record of each of the members of Groups I and II. A comparison between the

TABLE 7

Group I

SUBJECT	FINAL SCORE	NUMBER OF LAST ATTEMPTED	ARITHMETIC			OPPOSITES			ANALOGIES			TRUE—FALSE			INFORMATION			NUMBER SERIES			TOTAL OMITS	TOTAL ERRORS
			1*	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3		
1	146	160	17	3	1	35			34			21			35		9	18		1	3	11
2	147	162	17	2		36	1		35	1	3	21			35	2	4	18	1	1	7	8
3	150	154	17		2	34	1		34			20			33		1	16			1	3
4	99	110	12	1	1	24			24		1	14			24	6	1	12		1	7	4
5	99	126	13	1	3	28		2	27		3	17		1	27	2	8	14		4	3	21
6	129	147	16	3	2	32		2	32			19		1	32	2	5	16			5	10
7	116	139	15		4	31		1	30		3	18		1	30		10	15		2		21
8	117	130	15		1	28		3	28			17		1	28		3	14		1		9
9	144	151	17		2	33			33		1	20			32		3	16		1		7
10	107	127	14		2	28		2	27		3	17		1	27		7	14		2		17
11	96	118	13			26		4	26			15		2	26	3	7	12			3	13
12	123	129	14	1		28			28			17			28	2	2	14	1		4	2
13	113	137	15	1	1	30		5	18		1	18		1	30	5	3	14	1	1	7	12
14	153	177	19	8	1	39		1	38		1	23			39		5	19	5	2	13	10
Total..........	1,739	1,967	214	20	20	432	2	20	414	1	16	257		8	426	22	68	212	8	16	53	148
Average per person.........	124 2	140.5	15.3			30.9			29.6			16 9			30.4			15.1			3.8	10.6

* 1, number covered; 2, number omitted; 3, number wrong.

TABLE 8

Group II

SUBJECT	FINAL SCORE	NUMBER OF LAST ATTEMPTED	ARITHMETIC			OPPOSITES			ANALOGIES			TRUE—FALSE			INFORMATION			NUMBER SERIES			TOTAL OMITS	TOTAL ERRORS
			1*	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3		
1	85	101	11			22		1	22			13			22	8	2	11		4	8	7
2	111	118	13		2	26			26			15			26	3	2	12			3	4
3	109	126	13		3	28			27		2	17		1	27		7	14	1	2	1	15
4	113	147	16		11	32		3	32		1	19			32		7	16		9		31
5	164	177	19		2	39		1	38		1	23		1	39		5	19		1		11
6	137	146	16		3	32			31			19			32		4	16		2		9
7	132	150	17		2	33		1	33		4	19		1	32		7	16		1		16
8	107	111	13		2	24			24			14			24		1	12		1		4
9	105	139	15		1	31	7	3	30		1	18		1	30	13	4	15			20	10
10	104	118	13		3	26	1		26		1	15		1	26		5	12		2	1	12
11	138	148	17		5	32			32			19			32		2	16	1	2	1	9
12	133	155	17		3	34		3	34		3	20			34		6	16		4		19
13	139	150	17		1	33		1	33		1	19			32		6	16	1		1	9
14	115	120	13			26			26			16			26	1	2	13		2	1	4
15	108	110	12		1	24			24			14			24			12		1		2
16	119	148	17		3	32		6	32			19		1	32		11	16	1	1	1	22
17	134	145	16		2	32			31		1	19		1	31	3	2	16	1		4	6
18	100	106	11			24			22			14			23		4	12		2		6
19	154	164	18	1		36			35		1	21		1	36		3	18	2	1	3	6
20	158	176	19		6	39			38		1	23			38		10	19		1		18

21	133	165	18	1	3	36		6	36		3	21			36	1	10	18		3	2	25
22	101	124	13	2	3	27		5	26		1	17			27		1	14	1	5	3	15
23	128	144	15	1		32		1	30			19		2	31		9	16		2	1	14
24	105	134	15	2	4	29		5	29		2	18		1	29	1	4	14	1	3	4	19
25	113	123	13			27		1	26		1	17		1	27		4	13		1		8
Total..........	3,045	3,445	377	7	60	756	8	37	743		24	448		12	748	30	118	372	9	50	54	301
Average per person..........	121.8	137.8	15.1			30.2			28.3			17.9			29.9			14.9			2.2	12 0

* 1, Number covered; 2, number omitted; 3, number wrong.

reasons for omitting and the speed and accuracy scores of the subjects is, in many cases, illuminating.

THE MEASUREMENT OF "CAUTION"

Having established, through analysis, the fact that omissions were controlled (1) by the difficulty of the material, (2) by the form of the answer, and (3) by the attitude toward the directions, our next matter of interest was to see if this third factor, "attitude," could be measured.

Brown,[6] in his thesis, has attempted to measure a similar factor present in intelligence test performance. He has called his factor "caution." He has defined "caution" as a trait of human personality, as opposed to "rashness." A cautious person is one of the "careful," "conservative," type, while one who tends to be the opposite extreme is commonly called "rash" or "careless." The number of wrong answers made on the intelligence examination, were taken as indicative of the presence or absence of the "caution" factor. The greater the number of errors, the less "cautious" was the subject. (It should be recalled that his instructions did not forbid skipping.)

In discussing the situations in which caution is likely to be operative, Brown answers the following question. "Under what circumstances, does this factor (caution) if present in an individual, make its appearance?" His answer is:

> No well-defined situation may be described in answer to this question and regarded as a typical one. Generally speaking, lack of knowledge, either of facts or methods of procedure, coupled with other elements which make the situation an important one, such as the necessity for speed, the urgency of the occasion, and the character of the probable outcome, furnish a sufficient stimulus for the activity of the caution factor.

Thus, if the elements of the situation demand, in their handling, operations which are so simple for the subject that the content and technique, the speed, the urgency and "probable outcome" factors have the minimum of force, then the "caution" measure is vitiated, simply because the material for stimulating it is lacking. This is one variable which has to be controlled in measuring "caution" on an intelligence test. In an earlier paper[7] we have shown that different types of questions are of unequal difficulty for the same and for different groups. For example, arithmetic problems are more difficult for academic groups than analogies, and analogies are more difficult for policemen than arithmetic problems. Yet, both arithmetic problems and analogies are harder for policemen than for academic groups. Therefore, caution comparisons should be made only between similar groups—groups where the strength of the stimulus is comparable. Otherwise we are measuring mental acuity rather than caution.

---

[6] Brown, Wm. M., Character Traits as Factors in Intelligence Test Performance, Archives of Psychology, May, 1923, No. 6.

At the other extreme is the situation where the examples are so easy that few errors or omissions are found. Thus, 10 out of 25 subjects in Group II, made neither errors nor omissions on analogies. In the cases of these ten people, we do not know whether caution showed its influence in slowing up the answering process, or whether the questions were so simple and the answer so evident that caution had but little chance to operate. This is certainly a different "caution" setting than in the case of policemen, where 81 out of 81 persons made errors or omissions on analogies, and where the average percentage of errors was 35 per cent and of omissions, 28 per cent.

This is why we have limited our present study to Groups I and II where the final score, the speed, and the relative difficulty of the six types of questions are about the same.

Perhaps it is fitting at this time, to discuss in detail what we have meant by the terms "caution" and "cautious." In general, we have adhered to a behavioristic point of view and have considered persons cautious if their reactions in situations capable of stimulating caution, have been such as would be described as wary, heedful, and hesitant. But, in addition, we have taken cognizance, particularly in the introspective reports, of the central processes, which took place in the interval when the subject was making a choice between stimuli to react to and was determining the form of the reaction. Thus, we have been interested in discovering group differences in the manifestations of activities which would be described as cautious or rash, as well as the psychical processes which led individuals to make these particular motor reactions.

---

[7] Manson, Grace E., Group Differences in Intelligence Tests: The Relative Difficulty of the Types of Questions.

Webster's Collegiate Dictionary describes a cautious person as one who is "attentive to examine probable consequences of acts so to avoid danger." Now persons taking Test VI were placed in a setting, where it was necessary for them to become attentive to what consequences of their acts they considered the least dangerous. This conflict became acute on those questions where the correct answer was not immediately known.

In the main, subjects were attentive to the following three classes of stimuli, making those responses which had for them the least dangerous consequences.

1. Some subjects were attentive to the *instructions* not to omit and made errors knowingly, because they were instructed to guess.

2. Some subjects were attentive to the fact that *time* was passing and that it was desirable to answer as many questions as possible. These subjects omitted questions that were difficult and time consuming and guessed at others in order to cover as much ground as possible.

3. Some subjects were particularly attentive to the need for putting down only *correct* answers. They considered wrong answers as particularly undesirable and therefore worked slowly and checked answers, which repetition would lessen the number of examples covered. Difficult questions were frequently omitted in order that the subject might not injure his intellectual reputation for accuracy.

Which of these three types of persons shall we call cautious? We cannot see that there is any particular advantage in so delimiting the definition of the term cautious to exclude one type of reaction or the other. However, we do consider it of great importance to ascertain what mode of behavior, in relation to "caution," a subject is likely to choose when placed in a situation where a choice is possible; also to see if groups composed of persons with the same mental acuity but different occupational interests will make different choices.

In the earlier part of this report, we have shown that the members of Group I were more likely to omit on arithmetic questions than on other types of questions and that these omissions were made in order to gain time; whereas members of Group II were more likely to follow instructions, making errors rather than to omit.

But in both groups there were persons who did not conform to the general tendencies of the group. From the practical point of view of selection, these are the persons whom we should like to have differentiated immediately. It would be highly desirable if these persons could be distinguished by some objective scoring of their mental test records. This is a difficult task and we do not think that we have measured the trait "caution" with a high degree of reliability. Yet, we have attempted to do this, using the following method.

We have limited the measuring process to Group II because from both the introspections and the statistical analysis, it seemed as if the "caution of intellectual reputation" (desire not to commit one's self) and "caution of directions" (did not omit because told not to), were the more predominating ones in this group.

To obtain a criterion for such a scoring method, we asked the members of Group II to rate themselves on Bureau of Personnel Research Graphic Rating Scale, Form R-1. It was unfortunate that we had to use self-ratings as they are undoubtedly of low reliability, but members of the class were not well enough acquainted to make ratings of one another. Therefore, the correlations reported must be interpreted with the conditions of obtaining the criterion in mind. The steps on the graphic rating scale for caution were as follows:

Always acts on the spur of the moment	Impulsive. Always makes prompt decision	Shows moderate deliberation	Cautious, deliberate and considerate	Extremely wary and hesitant. Acts only after careful consideration

Scores were obtained by measuring the distance from the left hand end of the line to the point where the subject put his check. High numbers indicate a great deal of caution, low numbers, rashness.

Table 9 shows the correlation coefficients between the various caution indices and the caution self-ratings.

When these variables are combined in multiple correlation, the multiple correlation coefficient is raised to 0.42 ± 0.08. Although the zero order correlations are low, the directions which they take are interesting. Errors correlate negatively with self-ratings on caution: the more errors the less cautious. Omissions correlate positively with caution, speed correlates negatively with caution. Errors and omissions have a low negative intercorrelation. Speed and errors have a positive correlation. Speed and omissions have a low negative correlation. This latter intercorrelation is probably affected by two types of caution, the urge to skip difficult problems which would correlate positively with speed, and the urge not to put down incorrect answers even after consuming time in working them, which would lessen speed and result in a negative correlation. The latter tendency, in this case, seems to be the stronger one.

TABLE 9

		1	2	3	4
1	Caution rating				
2	Per cent of errors	−0.32			
3	Per cent of omissions	+0.25	+0.05		
4	Speed (number of last example attempted)	−0.22	+0.37	−0.22	

In order to put the measure of cautiousness into one number, we have formulated what we have called the "rashness" index, or index of "lack of caution." Since errors correlated negatively with caution ratings and omissions correlated positively, the ratio between these two gives a measure of "rashness" and "caution." The ratio form also better equates the difficulty of the material since the "rashness" index is the same for a person who has eight errors and two omissions as for one who has four errors and one omission.

Since the problems were probably more difficult for the former person than for the latter, this procedure corrects for difficulty, assuming that if the situation were of equal difficulty for each, the degree of caution would be the same. There is a certain fallacy in this hypothesis, in that caution may not increase with difficulty of the situation in the ratio that we are assuming. *In other words, we do not know the regression lines* for either difficulty of situation upon caution or caution upon difficulty of the situation. However, since we are measuring caution rather than intellectual acumen, it seems necessary to consider the caution variable when the caution stimuli are of equal strength.

In the "lack of caution" index, per cent of errors has been taken as the numerator and per cent of omissions as the denominator. The larger the index, the greater "rashness" or the less "caution."

This index correlated -0.32 with caution rating. If speed is added to the equation, the multiple correlation coefficient becomes 0.35. This is not quite as high as the multiple correlation using omissions, errors and speed as separate items.

Correlations were also found between self-ratings on other traits and "lack of caution" index.

"Lack of caution" index

Self-ratings on orderly	$r =$ 0.08
Self-ratings on submissive	$r =$ −0.01
Self-ratings on heedless	$r =$ −0.10
Self-ratings on wilful	$r =$ 0.34

Ratings on wilfulness correlated positively with "rashness" index. The other coefficients are too low to have any significance.

All of these measures taken from Test VI are complicated by the fact that some students were cautious of directions and wrote down wrong answers knowingly, simply because they felt forced to do so.

SUMMARY

We have studied objective records and introspective reports of two groups of persons making equivalent final scores on an intelligence test. By an examination of number of examples attempted, number of errors, number of omissions, and types of question having the greatest number of errors and omissions, we have found that omissions are controlled by form of question and answer, by difficulty of question, and by the reactions of the subject toward the urge to speed up, to disregard the directions not to omit, and to put down only correct answers. Verbal reports supplement and aid in interpretation of objective responses.

Groups I and II differ in respect to what may be called choice of "caution" stimuli. For caution may be exhibited by reacting to certain elements in the total test situation, with responses that make reactions to other "caution" stimuli impossible. For example, cautiousness of directions makes omissions impossible and incorrect answers probable. Cautiousness of accuracy, desire to have only correct answers, leads to omissions or few examples answered. Cautiousness of time, desire to finish the test, means likelihood of omitting or making errors on difficult and time-consuming problems. The choice between caution stimuli has to be made by the subject and the responses made are, as such, an expression of his personality.

On this basis, our results show that Group I subjects were more apt to respond to the speed factor than members of Group II, especially in those cases where the form of the question made guessing impossible. However, some subjects in Group I responded more actively to the request not to follow directions. Perhaps this tendency not to follow directions, when the subject foresees that more favorable results may be gained by omitting, could be called initiative, and the tendency not to follow directions and to omit, when the subject fears to put down the incorrect answer, lack of initiative or timidity.

As groups, we may then say that this factor of initiative was more often present in Group I, the business group, than in Group II.

An effort was made to measure the several caution reactions found in Group II. Self-ratings on "caution" were taken as the *criterion. Per cent of omissions correlated 0.25 with caution,* speed −0.22 with omissions, thus showing that the subjects making numerous omissions were inclined to judge themselves as cautious, and they were also inclined to be slow. On the other hand, persons who made errors were inclined to judge themselves as not so cautious, (correlation between percentage of errors and caution −0.32). The persons making errors also were more inclined to speed, the correlation between speed and errors +0.37. When a combination of the three items, speed, omissions, and errors, is made by multiple correlation, the coefficient of correlation increases to 0.42 ± 0.08.

The regression equation here given is probably of little value for other groups, due to the questionable reliability of the criterion as well as the internal consistency of Group II. Groups prone to omit for speed would require a different weighting. However, a detailed examination of errors and omissions on Test VI, supplemented by introspective reports of the subjects in respect to reasons for omission, will give insight into their personalities which is not ascertained through total scores alone.

PERSONALITY TESTS: WHITE AND NEGRO ADOLESCENTS[1]

DAGNY SUNNE

Newcomb College, Tulane University

The investigation here reported has been directed toward three questions: (1) Is there need of specific norms for adolescents, because these immature boys and girls give different forms of reactions from those given by adults to personality tests? (2) How do the results of such tests as the Pressey X-O tests, the Carnegie test IX (the Ream adaptation of the Downey Will-Temperament Scale), and the Woodworth-Mathews questionnaire agree when they are applied to adolescents? (3) What do such tests indicate as to similarity or diversity between white and negro boys and girls?

The Pressey X-O Scales[2] for measuring emotions were used with both children and adults. Form A designed for adults, consists of four tests attempting an inventory of what the individual finds are for him, (I) unpleasant associations, (II) common associations, (III) wrong acts, (IV) causes of worry and anxiety. Form B, adapted to children, is devised to ascertain what they feel to be (I) wrong, (II) causes of worry and anxiety, (III) interesting. As the individual tested must mark all the words that affect him unpleasantly or as wrong or interesting, and also circle the most unpleasant, worrisome, or interesting of the five words in each line, it may be possible to ascertain both the spread of emotionality and the preponderant emotional trends in regard to several hundred affective situations.

---

[1] This study was aided by a grant from the Commonwealth Fund.

[2] Pressey S. L., A Group Scale for Measuring Emotions, Jour. Abnormal Psychol., XVI, pp. 55-64.

Results from Form A have been obtained from 157 college students and 80 summer school students, all women. The latter group is somewhat less advanced scholastically but more mature than the former. The two groups have been kept distinct as they show some marked contrasts in responses to the tests. In test I the 125 words have been selected, as being associated with (*a*) sex, (*b*) disgust, (*c*) fear, or (*d*) suspicion. Some of the words in the first group seemed to have no sex-connotation for the students, as was ascertained indirectly through various association experiments after the tests had been taken. In a few cases the jokers may have distinctly unpleasant associations, as for example the word "fried." Bearing these reservations in mind, we find that the group of teachers circled about an equal proportion (27 per cent) of the words listed as "sex," "disgust," and "suspicion," and put the remaining 19 per cent in the "fear" list. Over 41 per cent of the words encircled by the student group were in the "disgust" list; 24, 18, and 16 per cent were circled respectively in the "sex," "fear," and "suspicion" lists.

The disagreement of the two groups as to particular words encircled is striking. Nearly every word in the "disgust" list seemed more unpleasant to a greater number of students than of teachers, and such terms as pervert, meanness, immoral, giggle, seemed much more offensive to the teachers than to the students. The norm was the modal word in every line except the thirteenth, where meanness proved equally as unpleasant as suck (the norm) because the teachers found the former most unpleasant, as opposed to the majority of the students.

In test II the modal words correspond to Pressey's norms with four exceptions as shown below:

5. Sleep;	worry (norm)	fright (modal)
11. dream;	beautiful (norm)	*floating (modal)*
23. sickness;	children (norm)	worry (modal)
25. death;	water (norm)	hopeless (modal)

In each case except the last, the teachers conform to the norms, but not the students. Only two words (sew and sale) are not circled at least once, hence individual variations are very noticeable and at times symptomatic of difficult adjustments.

Some interesting variations from the norms appear in test III. Here the college students are somewhat nearer the norms than the teachers. Both reckless and fussy are marked as worse than careless, the teachers especially emphasizing reckless; dishonest is circled three times as often as illegitimate, as is also debt as compared with chewing; stingy and meddling, are equally wrong according to the students, but meddling is three and one-half times worse in the teachers' estimations; cruel is much worse than street-walker, swindler and mistress than a dope-fiend; persecute and nagging, are about equally bad, though the students stress the latter. Three of the words are not encircled as blameworthy by anyone: dancing, tired, and church-going.

The modal words for test IV differed from the norms in 9 out of the 25 lines; in the case of 4 of these the majority of teachers encircled some other word and most of the students made other choices in another 4. "Money" was an insignificant third in degree of worries in both groups, as compared with failure and sickness. The other 8 modal words as compared with the Pressey norms were: conscience vs. clothes, falling vs. religion, cancer vs. insult, pain vs. blues, bashfulness vs. business, windstorms vs. syphilis, confusion vs. teacher, death vs. awkwardness.

If classified into the five groups suggested by Pressey, the words encircled by the student and teachers were grouped as follows:

	SUSPICIOUS	JUMPY	SELF-CONSCIOUS	MELANCHOLIC	HYPERCHONDRIACAL
	per cent	per cent	per cent	per cent	per cent
Students	15	18	18	26	24
Teachers	16	20	13	27	24

It is interesting to note that in both groups about 50 per cent of the encircled words belong in the melancholic and hyperchondriacal lists, and that the proportion of the self-conscious words selected is smaller for the more mature group.

As the modal words are found in different combinations in Forms A and B, it is difficult to make any reliable comparison for the two corresponding tests between children and adults. Both the women and the young girls agree in stressing the wrongness of ignorance, dirty, smutty, divorce, dishonesty (swiping), immodesty, and recklessness; but adults put greater emphasis on meddling, snobbishness and thoughtlessness; the children on smoking, worry and toughness. As to worries, both adults and children agree in finding failure and accidents, tuberculosis, fire, death, insanity (nervousness, nervous-breakdown) and forgetfulness the most serious. Religion, self-consciousness, longings are of much less consequence to the children than to the adults.

TABLE I

W is used to indicate the white children and N the negro.

TEST I (WRONG)	TEST II (WORRIES)	TEST III (INTERESTS)
1. Smoking	1. Forgetting (W)	1. Vaudeville (N)
3. Ignorance	4. Accidents and failure	2. Eating (N)
4. Fussing (N)	5. Disease	5. Studying (N)
5. Talking-back	6. Fire	6. Acrobats (W), minstrels (N)
6. Slang (W)	7. Tuberculosis	8. Hymns (N)
7. Bluffing	8. Death	9. Good boys
8. Kidnapping	10. Insanity (W)	10. Card parties and pageants (W)
9. Meddling	11. Storms (W)	12. Pop
10. Slow (N)	12. Suffocating	13. Typewriting
11. Faultfinding (N)	13. Poison	14. Actors (W)
13. Butting-in	14. Giggling	15. Gallahad (W)
14. Disgrace	15. Nightmare	16. Prayer (N)
15. Grumbling	16. Smoking (W), teasing (N)	17. Drawing (W), history (N)
17. absentminded (W), sly (N)	18. Drowning (W)	19. Traveling (W)
21. Bullying	19. Gun	20. Teachers and soldiers
23. Tobacco	23. Knives	21. Poems (N)
25. Borrowing	25. Rough	22. Ferris-wheel
		23. Bankers (W)
		24. Circus (W), cities (N)
		25. Joyriding.

A small number of teachers tried Form B and their modal words agreed with Pressey's norms in all but five lines, and in three of these they agreed with the children in their belief that smoking, bullying and borrowing are worse than begging, yellowness and flunking respectively. It may be of interest to note that the majority of children believe (as against the adults and Pressey's norms) that

Ignorance is worse than stinginess
Talking-back is worse than extravagance or boasting
Slang is worse than stubbornness
Kidnapping is worse than war
Tough is worse than cad
Butting-in is worse than cribbing
Disgrace is worse than immodesty
Grumbling is worse than snobbishness
Absentmindedness is worse than sham

Form B was given to 232 white and 282 negro children in the seventh and eighth grades. The agreement of a majority of these children with the Pressey norms for the modal words is not close, as out of 25 words in each test the modal words correspond to the norms as follows:

	I	II	III
White	12	0	11
Negro	9	13	10

The modal words for these two groups of children in the lines where they differ from the norms are shown in table 1.

As the white and negro children chose the same modal word in test I in all lines except six and in test II in all but seven, but in test III agreed only on eleven, there is evidently a greater consensus of opinion about wrongs and worries, as suggested by these tests, than about likes and interests.

It is interesting to note how evident is the effect of school training and social restriction in these choices. For example, the reason for the overwhelming vote against smoking, talking-back and slang is obvious. These children on the whole believe that fibbing and recklessness, betting and grumbling, absentmindedness and slyness, borrowing and teasing, divorce and faultfinding are about equally wrong; and that cowardice is much worse than temper, butting-in than stupidity, a strike than a trust, bullying than yellowness, tobacco than bribery, pool, or craps. The negro children emphasize that fussing, slowness, faultfinding, slyness, temper, kidding, fighting, quarreling, and teasing are wrong much more than do the white children. *Both groups find weakness, meanness,* nagging, delay, blarney, spending, hazing, queerness, neglect, prizefighting, dancing, unions, trusts, pull, and speculating least wrong. The difference between the modals of the boys and of the girls is very slight; the girls stress flirting, grumbling, and quarreling a little more, and the boys meddling, fighting, betting and bragging.

In test II, the modal words are not as clearly defined as in test I (especially in lines, 10, 14, 19, 21, 22, 23, and for the negro children also 15, 16, 17, 25, where several words are circled with about equal frequency) perhaps because these words do not represent children's most pronounced worries, or are of nearly equal importance in their experience, or because worries are more subject to individual diversity than wrongs. Forgetting and school, failure and accidents, reciting and suffocating, poison and lessons, jealousy, smoking and teasing, drowning and lies, crying and family, roughness and flightiness, fire and germs, knives and disposition, seem to have nearly equal influence in causing worries. The least worrisome appear to be depression, clothes, looks, self-consciousness, moodiness, queerness, politics, and wit. Insanity and flightiness are circled more often by the white children than by the negro, religion, soul, God, unfairness, and homeliness more by the negro. The sex differences seem insignificant. It is of interest that the worry least often circled in line 3 is clothes for the girls and looks for the boys. Do such results confirm a limitation of this type of test, as suggested by other investigators, that very unreliable answers are obtained from introverted, egocentric individuals? May it not be true of these adolescents that many of them are on guard against divulging their feeling at points where they are most sensitive?

Test III shows the greatest differences between negro and white children and also between the girls and boys. The two racial groups agree better in what they are least interested in, than in their greatest likes and interests. Fortune-telling, Raphael, day-dreaming, smoking, rough boys, palmistry, onions, loafing, Mowgli, deacons, crowds, and arguing are least often circled by both groups. Both also prefer athletic to pretty girls, minstrels and acrobats to talking, good and handsome boys to leaders (except the negro boys who are more interested in leaders than in handsome boys). But the negro children prefer hymns to jazz, pageants to card parties, musicians and artists to actors and aviators, poems to books, history to drawing, studying to dancing, and like babies and children better than do the white children. The negro boys are in better agreement with the girls as to chief interests than are the white boys and girls. Some of the more significant sex preferences are shown in table 2:

TABLE 2

WHITE		NEGRO	
Boys	Girls	Boys	Girls
Boating	Vaudeville	Boating	Vaudeville
Camping	Hiking	Napoleon	Edison
Napoleon	Beethoven	Saxophone	Singing
Pretty and athletic girls	Athletic girls	Pageants	Socials
		Sports (Babies*)	Babies (Sports*)
Reading	Dancing	Businessmen and soldiers	Teachers and nurses
Bands	Singing		Cities
Aviators	Actors	Circus	
Soldiers	Teachers		
Books	Newspapers		
Machines	Ferris-wheels		
Fishing	Joyriding		

* Second choice.

Besides special emotional trends, the Pressey X-O tests may also indicate the amount of emotionality by the total number of words underlined in each test. There are no constant differences between the seventh and eighth grades in the case of these children either in totals or deviations, perhaps because the numbers are small. A comparison with Pressey's norms for Form A (table 3) shows that percentiles of these college students and teachers agree with the norms in tests II and III, though there is tendency on part of the students to give fewer associations in test II, and on test III the teachers tend to find more "wrongs" and the students somewhat fewer. On tests I and IV, these averages are higher than the norms, more especially in the number of unpleasant associations. The deviations from the norms correspond very closely to those given by Pressey. For Form B, the mean deviations agree very closely with those given by Pressey for high school freshmen in the case of both white and colored children. There is a slight tendency for the deviations to be higher in the case of the negro children in "wrongs" than in "worries" and "interests." The total number of associations, both medians and means, vary so widely from Pressey's norms for tests II and III, that it is peculiar that the median for test I is so close to the norm. It is significant that the negro totals for "wrongs" is so much lower than for "interests" and that the opposite is true for the white totals. As can be seen from the percentile scores (table 3) the total scores for the groups tested agree with the Pressey norms only in regard to what is felt to be wrong and blameworthy, but vary considerably from the norms in the other tests. Where comparison is possible, the results for the children and adults agree closely except that the extreme scores of the children are greater. The scores of the two adult groups agree best in regard to worries (test IV) and differ about equally for what is wrong and unpleasant (tests III and I), the students averaging lower scores in both these tests.

To investigate the dynamic side of personality, Ream's[3] adaptation of the Downey Will-Temperament tests was tried after it was found impossible to use Miss Downey's individual form on account of the time required. In spite of the dubious interpretation of some of these tests, others seem to have pro-

---

[3] Ream, M. J., Group Will-Temperament Tests, Jour. Educat. Psychol., XIII, 1922, pp. 7-16.

TABLE 3

Pressey X-O Tests

Form A (totals)

	NUMBER	Test III (wrongs)				Test IV (worries)				Test I (unpleasant)				Test II			
		25 PERCENTILE	MEDIAN	75 PERCENTILE	HIGHEST SCORE	25 PERCENTILE	MEDIAN	75 PERCENTILE	HIGHEST SCORE	25 PERCENTILE	MEDIAN	75 PERCENTILE	HIGHEST SCORE	25 PERCENTILE	MEDIAN	75 PERCENTILE	HIGHEST SCORE
Pressey norms		60	73	86		33	46	55		27	41	62		41	55	70	
Students	157	61	69	77	101	42	50	61	119	42	49	59	101	41	49	61	102
Teachers	80	70	77	86	101	39	50	66	94	49	60	72	85	39	54	67	91
Total	237	63	72	80		41	50	63		43	53	68		39	50	62	

Form B (totals)

		Test I (wrongs)				Test II (worries)				Test III (interests)			
Pressey norms (High School freshmen)			70				24				40		
White													
Girls	102	59	74	93	123	41	59	67	123	45	56	69	109
Boys	130	61	75	93	122	31	41	60	121	46	59	80	112
Total	232	60	74	93		36	45	65		45	57	73	
Negro													
Girls	182	52	65	82	124	38	50	70	125	56	72	89	124
Boys	110	49	66	86	120	39	49	66	121	59	79	92	125
Total	292	51	65	85		38	50	69		57	72	90	

Form A (average deviations)

		Test III	Test IV	Test I	Test II
Pressey norms		13	15	11	10
Students	157	13	16	12	10
Teachers	80	12	16	10	10
Total	237	13	16	11	10

Form B (average deviations)

		Test I	Test II	Test III	
Pressey norms		17	18	18	
White:					
Girls	102	16	17	19	
Boys	130	17	18	17	
Total	232	17	18	18	
Negro:					
Girls	182	18	17	17	
Boys	110	18	17	17	
Total	292	18	17	17	

nounced value, so that it was deemed best to use the whole series, more especially as comparative measures might give information about some of the doubtful tests. In fact, the use of Miss Downey's tests with individual college students had convinced the writer of their value in the analysis of personality.

When the scores, grouped according to the Carnegie percentiles for adults, are compared (table 4), it is seen that the differences between boys and girls are small, except in tests XI and XII for the negroes, where the divergence is more than one percentile. As for the racial groups, the negro boys have somewhat lower averages in the first four and the last four tests, and equal or higher averages in the middle four than the white boys; between the two groups of girls, the differences are less, the negro girls having lower scores on tests I, II, and IX (speed of movement, freedom from load, coördination of impulses), and higher on tests VI and VIII (motor impulsion and freedom from self-consciousness). When the scores are combined, the negro children have lower averages on tests I, III, and IX (speed of movement, speed of objective

TABLE 4

Percentiles in test IX (according to Carnegie adult norms)

	NUMBER	1	2	3	4	5	6	7	8	9	10	11	12
Negro boys............	84	3.6	5.1	1.5	4.5	4.5	6.6	6.1	7.5	1.8	2.4	3.9	5.8
Negro girls............	199	4.2	5.2	1.8	4.4	3.8	6.7	6.2	7.0	2.1	1.9	2.5	7.0
Total...............	283	4.1	5.2	1.7	4.4	4.0	6.7	6.1	7.2	2.0	2.1	2.9	6.7
White boys.............	59	4.9	5.9	2.4	5.4	3.0	5.2	6.4	6.0	3.1	3.1	3.1	6.2
White girls.............	81	5.2	5.2	2.5	4.4	3.6	5.7	6.6	5.9	2.9	2.2	2.0	6.8
Total................	140	5.0	5.5	2.5	4.7	3.3	5.5	6.5	6.0	3.0	2.6	2.5	6.6
Students..............	150	8.2	6.1	5.8	4.7	5.2	6.2	7.3	2.6	7.1	6.5	4.5	6.8
Teachers..............	67	6.0	4.6	5.3	5.7	5.5	6.6	6.2	4.2	3.8	3.6	3.0	6.2
Total (adult).........	236	7.0	5.7	5.7	5.0	5.3	6.4	7.0	3.1	6.1	5.6	4.1	6.7

decision, coördination of impulses), but higher on tests VI and VIII than the white children. The year-groups are probably too small, for no constant tendencies are evident, except that of higher scores among the younger than among the older children on tests requiring speed. The averages of the adults for these same tests are higher by 3 percentiles or more in tests I, III, IX, and X, and as much lower in test VIII; and between 1 and 2 percentiles higher in tests V and XI.

In order to see whether more significant trends might appear if grouped according to Miss Downey's suggestions, the scores on tests I, II, III, and IV were combined, and VI, VII, VIII, and IX, X, XI and XII (test V being omitted as too liable to error both in scoring and interpretation), thus forming patterns A (the mobile, rapid-fire type), B (the aggressive), and C (the deliberative type).

TABLE 5

Reaction-pattern percentiles in text IX (according to Carnegie adult norms

	NUMBER	A 1–4	5	B 6–8	C 9–12	TOTAL
White boys	59	18.4	3.0	17.6	15.6	54.6
White girls	81	17.3	3.6	18.2	14.1	53.2
Negro boys	84	15.3	4.5	20.0	13.8	53.0
Negro girls	199	15.7	3.8	19.9	13.5	52.8
White total	140	17.8	3.3	18.0	14.7	53.8
Negro total	283	15.4	4.0	20.0	13.6	53.0
Students	169	24.9	5.2	16.1	24.9	71.1
Teachers	67	21.6	5.5	17.0	16.6	60.6
Adult total	236	24.0	5.3	16.4	22.5	68.3

According to this grouping (table 5), the white boys have a slight advantage over the girls in patterns A and B, while the negro boys and girls have almost the same averages. The average total scores show no difference between these groups. The adults have decidedly higher averages for patterns A and C but lower for pattern B.

Because individual variations on the tests are so great, another attempt at comparison was made by finding the groups in which the average of each individual tested was at or above the fiftieth percentile. The percentage of each of the groups

TABLE 5

Percentile table for adolescents

	FIRST QUARTILE	SECOND QUARTILE	THIRD QUARTILE	FOURTH QUARTILE
White children				
Speed of movement	35-48	49-54	55-62	63-84
Freedom from load	0.20-0.39	0.40-0.44	0.45-0.54	0.55-0.84
Speed of objective decision	2-14	15-17	18-22	23-36
Speed of subjective decision	1-16	17-20	21-25	26-36
Flexibility in disguise	0	1-2	3	4-6
Motor impulsion	0.03-1.69	1.70-1.85	1.80-2.09	2.10-2.99
Assurance in visual memory	0-12	13-18	10-22	23-31
Freedom from self-consciousness	0.05-0.64	0.05-0.79	0.80-1.05	1.06-5.60
Coordination of impulses	10-49	50-59	60-72	73-103
Motor inhibition	1-18	19-39	40-55	56-89
Care for detail	0-9	10-15	16-19	20-34
Perseveration	20-82	83-105	106-139	140-249
Negro children				
Speed of movement	19-43	44-50	51-60	01-120
Freedom from load	0.14-0.35	0.36-0.44	0.45-0.51	0.52-1.00
Speed of objective decision	1-9	10-13	14-17	18-36
Speed of subjective decision	2-13	14-20	21-24	25-26
Flexibility in disguise	0-1	2	3	4-0
Motor impulsion	1.20-1.82	1.83-2.01	2.02-2.25	2.26-4.25
Assurance in visual memory	0-11	12-16	17-22	23-31
Freedom from self-consciousness	0.05-0.53	0.54-0.69	0.70-0.89	0.90-3.09
Coordination of impulses	1-42	43-52	53-62	63-100
Motor inhibition	1-31	32-49	50-68	69-90
Care for detail	0-9	10-14	15-20	21-34
Perseveration	10-86	87-106	107-139	140-348

that were at or above the fiftieth percentile in the first reaction-pattern alone, only in the second and only in the third; in both first and second, both first and third, both second and third, and above that point in all three patterns are given below. The negro children were classed as twelve to fifteen-year-olds and sixteen-year-olds, because both boys and girls were thus about evenly divided and especially because there were some peculiar differences between the younger and older children (see table 7).

The percentages of those tested whose averages were below the median in all the reaction-patterns were: students 6 per cent; teachers 3 per cent; white boys 7 per cent; girls 5 per cent; negro boys 4 per cent; girls 6 per cent. According to the Carnegie norms for adults, the predominance of the B pattern with the children and of the ABC pattern with college students is equally conspicuous. The preëminence of the AB pattern with both children and teachers may also be worthy of comment, as well as the infrequent occurrence of the AC pattern among the negro and its relative importance among

TABLE 7

PERCENTAGES ACCORDING TO ADULT NORMS	PATTERNS						
	A	B	C	AB	AC	BC	ABC
White boys	7	42	2	16	11	7	15
Negro boys:							
12–15 years	2	53	0	22	0	11	0
16 years	5	60	0	21	0	8	5
Negro boys (total)	4	57	0	22	0	10	7
White girls	5	47	1	27	5	1	13
Negro girls:							
12–15 years	3	41	1	38	0	8	10
16 + years	2	67	1	21	1	7	2
Negro girls (total)	2	56	1	28	1	7	5
White total	6	45	2	23	8	4	14
Negro total	2	50	1	26	1	8	6
Students	2	2	2	3	19	4	67
Teachers	8	15	2	31	9	8	28
Total (adults)	3	6	2	10	17	5	56

the white subjects. The two most frequently occurring patterns among the children are the B and AB, among the students ABC and AC, among the teachers AB and ABC. The least common for all groups are the A and the C patterns, though among the teachers the A pattern is found as often as the AC and BC.

When a similar grouping of reaction-patterns is calculated from the norms of the white children (as seen in table 8) there is a conspicuous preponderance of A and B alone and combined with C in the case of negro boys; and equally significant are combinations of patterns rather than preponderance of a single pattern as regards the white boys. The negro girls show a similar lack of the A type as do the white children, and predominance of the BC-groupings, a tendency that is peculiarly evident among the older negro girls. The difference between the negro boys and the rest in percentage of the AB

TABLE 8

PERCENTAGES ACCORDING TO WHITE ADOLESCENT NORMS	BOYS						
	A	B	C	AB	AC	BC	ABC
White boys	10	4	8	12	23	27	18
Negro boys:							
12–15 years	25	19	11	6	14	19	6
16 + years	15	24	6	12	15	15	12
12–16 years	20	22	9	9	14	17	9
White girls:	8	8	13	12	17	16	26
Negro girls:							
12–15 years	6	11	12	12	22	16	20
16 + years	0	16	18	8	10	30	17
12–16 years	3	14	16	10	15	24	18
Boys and girls:							
White	9	6	11	12	20	20	23
Negro	8	16	14	10	15	22	16

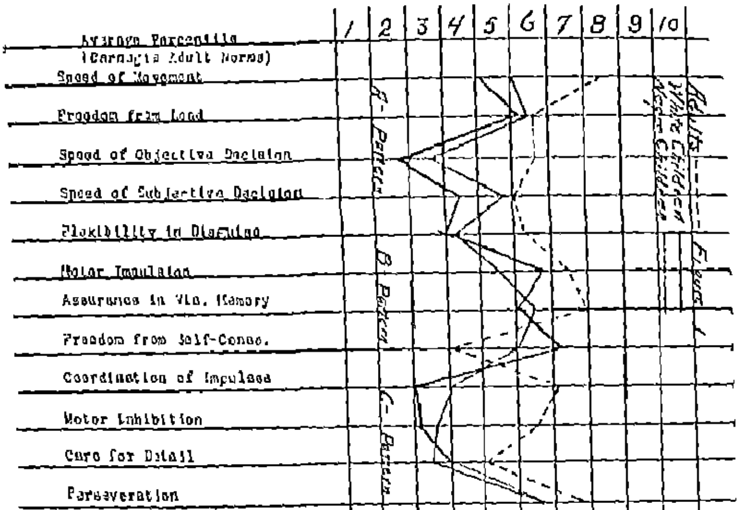
patterns is also great, 25 per cent more than the negro girls and 27 per cent more than the white children. The white children and the younger negro girls average about 38 per cent for the B, C, and BC groups, the negro boys about 48 per cent, but the older negro girls 64 per cent. Fewer girls than boys failed to reach the white median in some one of these patterns as the percentages of those below the medians were: white boys 12 per cent, girls, 6 per cent; negro boys 18 per cent; younger negro girls 5 per cent; older 9 per cent.

The temperamental patterns obtained by using the adult as compared with the adolescent norms are so divergent that there seems to be valid ground for further investigation of possible differences either in traits or in the norms that must be used for different ages, or in both. The predominance of the B and the AB patterns according to the adult standards (two-thirds of the negro and four-fifths of the white score are at or above the medians in these patterns) is as clear as the more diversified groupings according to the adolescents standards (less than one-fourth are found in the A and AB patterns). According to the white adolescent standards, about one-half of the negro boys, one-third of the negro girls, and one-fourth of the white children belong to the single-type patterns. *It is evident that the curves of the average percentiles (fig. 1) for the white and negro children are much* more similar than is either to the adult curve.

A comparison of these results with those of McFadden-Dashiell[4] are of interest. Both agree in finding the negro slower in movement, having the same or slightly greater inertia, greater motor impulsion, and the same interest in detail; they disagree in finding the negro having about the same volitional perseveration and coördination of impulses, and less motor inhibition. The latter trait is one in which there seems a distinct decrease with age among the adolescents tested, perhaps on account of the type of older children that are retarded according to school grades. It seems also that the

---

[4] Racial Differences as Measured by the Downey Will-Temperament Test, Jour. Applied Psychol., VII, 1923, pp. 30–53.



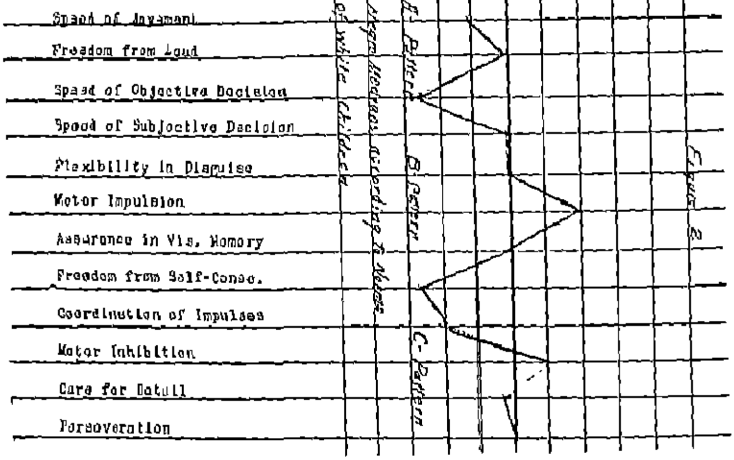


FIG. 1. Curves of average percentiles in Reams Downey Will-Temperament Test for adults----, white children——, negro children——.

FIG. 2. Negro children medians below or above heavy horizontal line which represents norms of white children.

younger negro children show more speed and less aggressiveness than the older. The degree of intelligence, amount of school training and maturity may all combine to produce such differences. A similar variation appears between the teachers and the college students.

The chief disagreement of the results from these adolescents with the McFadden-Dashiell findings is in regard to the mobile (A) and deliberate (C) patterns. Neither according to their own norms or those of the adults are the white children clearly superior in these two patterns. The negro boys have a distinct advantage in the A pattern, while the negro girls have exceedingly low percentages in this group. In the C pattern the negro children average lower in some of the tests and higher in others by both the adult and the white children norms. According to both adult and adolescent norms, these white children surpass in two types; the mobile combined with the deliberate and a combination in which all three patterns are above the median; but the negro somewhat in the aggressive type. In the other patterns, the two groups taken as a whole, show no distinct differences. As to variability, the results of these tests agree with the investigation just mentioned in finding the negro slightly more variable.

A recent comparison of these patterns with Jung's types[5] finds high scores in the A pattern corresponding to the extravert and high scores in the C pattern corresponding to the introvert type. The negro children have lower medians in the speed tests, but not less flexibility, greater interest in detail or perseveration than the white children, so there is no distinction as to introversion or extraversion in the middle ranges of scores. At the lower and higher extremes, the tendency is toward extraverted patterns on the part of the white children and introverted on part of the negro, according to the scores on the tests.

The revised form[6] of the Woodworth-Mathews questionnaire was tried by 81 white and 302 negro children, all in the eighth grade of the public schools. The number of unfavorable responses ranges from 3 to 25 for white boys, from 1 to 37 for white girls, 1 to 34 for negro boys, 3 to 38 for negro girls. The percentile distribution of unfavorable responses is as follows:

---

[5] Downey, J. E., Jung's "Psychological Types" and Will-Temperament Patterns, Jour. Abnormal Psychol., XVIII, 1924, p. 345.

[6] Mathews, E., A Study of Emotional Stability in Children, Jour. Delinquency, VIII, no. 1, p. 39.

	WHITE				NEGRO			
	Number	25 percentile	Median	75 percentile	Number	25 percentile	Median	75 percentile
Boys	48	7	10	14	95	8	13	18
Girls	33	10	14	19	207	12	17	22
All	81	7	12	17	302	11	16	21

This shows a greater frequency of such reactions on the part of girls than of boys and in the case of negro than of white children. Moreover, 8 per cent of the white girls (and 2 per cent of the negro girls) reach or exceed the highest score of the boys.

In regard to the responses to the 75 questions as compared with Miss Mathew's percentages for unselected boys and girls, the outstanding fact is general agreement. The slight variations may be of interest. These children give more unfavorable responses to four but fewer to twelve questions than did Miss Mathews' group. The negro children, in addition to similar differences as to white children, have higher percentages for seven additional questions and lower for two. Social restrictions offer some explanations for variations to questions concerning fear of dark, nobody understands, other people like you less, what to do next, and seeing red when angry. Though there is higher frequency in case of some questions, there is also much more constricted range of unfavorable answers, and smaller average scores, so the results can hardly be used as an argument for greater emotionality on the part of these groups of children than was found in the case of Miss Mathews' unselected boys and girls.

In regard to the relation of age to the number of unfavorable responses, the numbers are too small to give more than an indication of trend, as can be seen from the following means:

	WHITE				NEGRO			
	Boys		Girls		Boys		Girls	
12–13–14 years old....	(27)	10.6	(38)	15.9	(33)	14.2	(48)	17.1
15–16 years old.......	(6)	8.5	(9)	10.0	(62)	12.3	(159)	17.8

There seems to be a decrease with age, except in case of the negro girls, whose average remains almost constant at every age. The difference is also less between the younger of the white and negro girls than it is between the similar groups of boys.

The questions which elicit the greatest difference in responses between the white and negro children are the following, for which the negro percentage exceed the white by about 13 per cent or more:

8. Uneasy to cross a bridge
9. Going into tunnel or subway
11. Afraid during thunder storm
13. Afraid of dark
22. Habit of stuttering
28. Food makes you sick
40. Nobody quite understands you
46. Things get misty before your eyes
48. Pains in any part of the body
54. Eyes often pain you
64. Uneasy in a small room with the door shut

The white children have higher percentages for the following questions:

19. Do you dream of robbers?
20. Have the same dream over and over?
44. Idea that you are an adopted child
57. Have you often fainted away?
74. Feel you were very wicked?

The most popular questions (total percentage of unfavorable answers about 40 per cent or over) for both white and negro children proved to be:

11. Afraid of thunder storm
15. Afraid of noises in the night
17. Dream about your family
18. Dream about people being dead
19. Dream of robbers
29. Food so disgusting that you cannot eat it
32. Feelings hurt so badly that you cry
40. Things seem to get misty before the eyes
67. Great fear of fire

The percentage of unfavorable replies of both white and negro girls is considerably higher than that of the boys to the questions about:

9. Going into tunnel or subway
10. Afraid of water
11. Afraid during a thunder storm
14. Frightened in the middle of the night
15. Afraid of noises in the night
18. Dream about people being dead
19. Food that makes you sick
32. Feelings hurt so badly that you cry
64. Uneasy in a small room with doors shut

But the boys have the higher percentages for 5 (want to run away from home) and 25 (habit of twitching head, neck or shoulders).

A rough attempt was made to find the relationship between these three personality tests by comparing the number of scores above and below the group median of each test with every other. Of the white children who answered the Woodworth-Mathews questionnaire, 72 had also taken Pressey's X-O Form B. Those whose scores on the questionnaire were below the median and therefore by this test perhaps the more stable emotionally, had on the average smaller totals on tests I and II (wrongs and worries) and larger on test III (interests) than those with a higher frequency of unfavorable answers to the questionnaire, and their deviations averaged higher on tests I and II and lower on test III. Of the 62 negro children who answered both the questionnaire and the Pressey X-O, those also averaged higher on the deviations in test I and II and lower on test III, whose scores were below the median on the questionnaire, but their totals were a little lower on test III, and about equal on test I and II. Since the deviations on test II (worries) are higher in the case of both white and negro children who give fewer unfavorable replies on the questionnaire, the two forms of testing apparently do not agree. However, if the modal words of this same group of children is used as a criterion for test II, the results agree better with those of the questionnaire.

The scores on the Will-Temperament Test IX of those above and below median on the questionnaire show that the children both white and negro who have the fewer unfavorable responses are on the whole less speedy and exhibit less load; but the white children also score lower on coördination of impulses and motor inhibition, while the negro children show the opposite tendency except the older girls, who also score lower on motor inhibition. As about 200 negro children took both tests, and only 81 white children, it may be that more weight should be given to the results of the former. In the case of the negro children, those who have the fewest unfavorable replies on the questionnaire show a distinct tendency toward less inertia, better coördination of impulses and less perseveration and somewhat less speed than those who give the greater number of such responses.

A similar comparison of the three patterns of test IX with the questionnaire gives the following results: The white children who have the lowest scores on the questionnaire tend to have higher scores on patterns B and C and the negro children only on pattern C. In both cases, there is no difference in the percentages of those above and below the median on the questionnaire as regards pattern A, the mobile type. Both white and negro children who score higher on the deliberative pattern of test IX, also give fewer unfavorable responses to the questionnaire; the aggressive pattern B of test IX does not differentiate the negro, but is the most differentiating in the case of the white children.

A comparison of those above and below their group medians on X-O Form B and on test IX shows 20 per cent or more difference in the percentages as follows:

According to total scores. *a.* X-O I (Wrongs). Higher scores on coördination of impulses and lower totals for both negro and white.

b. X-O II (Worries). Higher scores on speed of objective and subjective decision correspond with higher totals on test II on the part of the negro children; no distinct differentiation in the case of the white children.

c. X-O III. Higher scores on assurance and higher totals on X-O III (interests) correspond for the white children and the negro boys. In the case of the negro children, higher scores on the speed tests of the Will-Temperament scale and higher totals on the Pressey Test occur together.

d. Higher scores on speed of objective decision and less deviation from the modal words of X-O I for both white and negro, but higher scores on freedom from load correspond with greater deviation for the white and less deviation for the negro; higher scores in perseveration with less deviation on the part of the white children.

e. Higher scores on care of detail and on perseveration occur together with greater deviations in X-O II.

f. Greater deviation from the Interest norms (test III) corresponds in the case of the white children with greater motor inhibition and in the case of the negro with greater speed of subjective decision.

IN CONCLUSION

The three personality tests used in this investigation show some significant agreements between parts of the tests. There are discrepancies for different age, sex, and race groups that may be partly due to unreliable norms. This is especially noticeable with reference to variations from the norms for total number of words crossed out in the Pressey X-O, Form B. Individual differences are, however, so striking that all three tests are valuable for diagnosis without regard to the normals.

The results give grounds for believing that there is need of investigating norms for different stages of maturity. An attempt at obtaining adolescent norms for the Carnegie Test IX is shown in table 6. By using these percentile scores instead of the Carnegie norms, the profiles of the children of both racial groups are very much alike and markedly different from those of the adults tested. Similar relationships are shown as to variability. The standard deviations vary from 0.9 to 9.7 for the twelve tests, but there is no significant difference on any of the tests between the racial groups examined in this respect. The highest standard deviations (almost equal for boys and girls, white and negro) occur in the tests for speed of objective and subjective decision, care of detail, and assurance in visual memory.

The Pressey tests elicited greater variations between the racial groups than the other two tests and seem to reflect differences in social restrictions and taboos. They also point to the importance of distinguishing between different forms of emotional responses, as the spread of emotionality may be more or less restricted in respect to anger, grief, fear, interests, in very divergent degrees in the same person, and perhaps for different groups. At any rate, the individual who has the greatest total on one of the Pressey tests may have a very low total on another. So also different groups show marked variations in totals while the deviations for each test remain relatively constant.

The differences between the racial groups examined can not be summed up in such generalizations as more or less emotional, energetic or unstable, or even by total scores on the tests. So many of the variations found may be due partly to lack of speed in writing and reading that very precise conclusions are out of place. Educational tests given to these same children have resulted in much lower scores in these subjects for the negro than for the white pupils. As is true of other forms of measurement, the investigator is impressed with the need of repeating the personality tests for several years with the same groups and combining them with studies of well-defined habits, in order to ascertain the significance of the results and obtain reliable norms. Thus it may be possible to discover what the trait investigated may mean in different situations, for these as well as other personality tests show that fruitful comparisons are dependent on varying degrees of these complex characteristics rather on the possession or non-possession of the traits. It may be added that the tests used have proved of considerable value in diagnosing difficult school and social adjustments. It is the accurate determination of these special situations that must await further study.

PERSONALITY AND VOCATIONAL ACHIEVEMENT

PART II

GEORGE C. BRANDENBURG
Purdue University

(Continued from June 1925)

A senior having completed a four-year record which was uniformly of such a nature as not to reflect great credit on him was heard to remark, "Yes, I admit that my record here is nothing to be proud of, but wait till I get out in real life and I'll show these professors that they have me sized up all wrong." And the instructor himself knows that college life after all puts the student to many tests and makes of him numerous requirements which often seem to be quite irrelevant to the later life and activities of the student. He knows all too well that the student who has been branded as a failure in his college work has ofttimes caused his former college instructor considerable embarrassment by his own subsequent success in one field or another. In fact, there is an occasional instance in which a student has been granted permanent leave of absence from the college because of scholastic delinquency who has later made a marked success in some useful vocation.

Again in various walks of life how often one hears the statement, "Well, I've been making a failure so far, but I've turned over a new leaf and from now on . . . ." and so on. But just how often does an individual turn over a new leaf and keep it turned? How far does the student's record in college predict his performance in life? To what degree are those personal traits by which a young person is known to his associates permanent and how will they affect his success in the world of affairs? The chief difficulty in answering such questions lies in the absence of any generally accepted definition of the term success. It is likely that in the popular mind the word, *success*, is practically synonomous with the word, *income*, and it is quite useless for idealists to point out that one's income is never a measure of its possessor's happiness or usefulness to society so long as most of the very persons who are engaged in pointing out this fact are nevertheless energetically devoting themselves in one way or another to the accumulation of wordly goods. Success is sometimes defined as the attainment of a worthy goal, but the difficulty with this notion of success is in the impossibility of agreement on what really constitutes a worthy goal. Certainly the accumulation of money may be a most worthy goal if the money is to be used in the improvement of the general health of mankind. Under certain circumstances almost any goal may be a worthy one. In the absence of an adequate definition of success or of the possibility of devising one, we shall find it preferable in our experimental work to indicate merely as specifically as possible the attainments or achievements of the various individuals concerned and the relationships of these achievements to other significant facts and conditions.

Vocational psychology is interested not only in methods of human analysis in order to aid the individual in the proper choice of a vocation but also in the determination of the human factors which are likely to be conducive to success in a given calling. For once such determinations are possible even within reasonable limits, society will be vastly more able to organize efficiently and satisfactorily its human forces; and the individual through the possibility of more intelligent planning of his life endeavors will be able to lead a happier, more useful, and more contented life. Moreover, through the determination of the most important personal traits for success in any given vocation it should be possible for education to organize its program to the greatest possible advantage of the individual and of society as a whole.

An attempt has been made by the author to determine what factors may be discovered in the young man's character or personality which would suggest his ultimate vocational selection and at the same time yield valuable information concerning the traits essential to success in the various vocations which engineering college graduates enter. It should be said that the methods and plans used were not selected because they were believed to be the most adequate or scientific, but because they were apparently the only ones possible under the circumstances and did seem to contain promise of interesting scientific adventure for all concerned. Consequently, following up the rating experiment described in the first part of this chapter, five and one-half years after the graduation of group I letters of inquiry were sent out to the members of both groups I and II, requesting detailed information concerning their activities along the following lines: (1) Nature of work, (2) writings, inventions and special honors, (3) earned income, (4) miscellaneous activities. The letter requested that specific information be given along these lines for each year since the student's graduation. Responses were received from 21 of the 29 men in group I and from 15 of the 25 in group II. In the remainder of the discussion we shall refer to these two groups as groups Ia and IIa.

VOCATIONAL DISTRIBUTION OF THE MEN

Presumably there are certain characteristics of an individual's personality or of his college performance which would indicate in what general field of activity he will eventually find himself. For unless he does materially modify his typical and habitual modes of thinking and acting after his college days are past, we should expect him to shift about in his vocational activities until he finds a kind of work in which he can express himself in ways which he has already found best suited to his nature and consequently most satisfying. If he were entirely free from chance influences we may be very sure that this is exactly what would happen; but during the first five years after graduation, at least, apparently these accidental or chance factors play havoc with our conclusions or predictions as to the type of work in which individuals will engage. For example, we find five of our group engaged in teaching despite the fact that not one of the men manifested any interest in this vocation during his college days. We may suppose that these men, being unable to obtain satisfactory positions in the fields for which they had specifically prepared, are simply marking time in the schoolroom until opportunities in their chosen fields open up. One man wrote that he was "just picking up a little ready money teaching while business in my line is a little dull."

The vocations represented by the two groups of men and the number engaged in each at the time of the inquiry are as follows:

Sales engineering	11
Telephone engineering	5
Teaching	5
Design work	3
Engineering research	2
Construction	2
Railway engineering	2
Estimating and appraising	1
Radio work	1
Bookkeeping and office management	1
Clerical work	1
Power engineering	1
Law	1

Twenty-eight of the men or a little over 75 per cent are in some kind of engineering work although in some cases, not the kind in which they expected or prefer to be, as their replies indicate. During the next five or ten years there will be considerable shifting about so that, as suggested above, most of the men will eventually get into lines of work which they prefer and for which they are by nature and training best fitted. There are naturally a few individuals who will discover that the work for which they have prepared and which they prefer is not at all suited to their abilities. It is clear that there is represented in this period of vocational adjustment a serious loss of time and effort to the individual and to society. Of course, the loss is small in such fields as engineering when compared with that in less highly specialized vocations.

As far as we can judge from a comparison of the men in the various fields there are no consistent or significant differences in their personalities, mental abilities, or scholastic records to which their present vocational choices can be ascribed. It is hardly to be expected that such differences would be very apparent in such a short time; but we confidently expect that significant differences will appear with a longer period of observation when the men have had opportunity to become somewhat better located vocationally.

TABLE 7

Showing rankings of the men in group Ia in (1) personality, (2) general intelligence, (3) school grades, (4) income, five years after graduation, together with number of writings, inventions, etc., by each

INDIVIDUALS	PERSONALITY	GENERAL INTELLIGENCE	SCHOOL GRADES	INCOME	INVENTIONS, WRITINGS
A	1	2	2	1	
B	2	4	4	8	
C	3	17.5	7	5	
D	4	14	11	4	1
E	5	1	1	6	4
F	6	5.5	5	10	3
G	7	9.5	18.5	15.5	
H	8	12.5	6	2.5	3
I	9	7	12	13	1
J	10	16	14	7	
K	11	5.5	8	15.5	
L	12	11	21	2.5	
M	13	No test	10	12	
N	14	8	9	11	
O	15	12.5	16.5	9	
P	16	9.5	3	21	1
Q	17	3	13	19	
R	18	19	18.5	20	
S	19	20	15	17.5	
T	20	17.5	20	14	
U	21	15	16.5	17.5	

PERSONALITY AND INCOME

Tables 7 and 8 have been prepared to show the detailed rankings of the individuals in the two groups with reference to personality, as determined by the rating experiment, general intelligence, as determined by the Army Intelligence test, school grades, as determined by the composite of the grades for the entire four-year college course, income, and writings and inventions. Since we have already discussed in the preceding section the relationships between personality, general intelligence, and school grades, we shall hereafter be chiefly concerned with personality, general intelligence, and certain phases of vocational achievement.

TABLE 8

Showing the rankings of the men in group IIa in (1) personality, (2) general intelligence, (3) school grades, (4) income five years after graduation, together with number of writings, inventions, etc. by each

INDIVIDUALS	PERSONALITY	GENERAL INTELLIGENCE	SCHOOL GRADES	INCOME	WRITINGS, INVENTIONS
A	1	8	1	3	5
B	2	No test	5	1	
C	3	1	2	8	1
D	4	5	4	11	
E	5	3	12	5	
F	6	12	10	12	
G	7	4	6	6	1
H	8	9	9	9	
I	9	11	7	2	
J	10	13	15	7	
K	11	14	11	4	
L	12	10	13	15	
M	13	2	3	13	2
N	14	7	14	10	
O	15	6	8	14	

Whatever may be said to the detriment of financial success as a measure of success in general it will readily be admitted that it does constitute at least one measuring rod for success and that it is very concrete and, therefore, very convenient to use as a basis of comparison among individuals. Further it may be safely assumed that most college graduates of the type we are here considering, that is, men who do not come from very wealthy homes, will be concerned chiefly in getting ahead financially during the first few years after graduation. We shall, therefore, assume that the relative incomes of these men make at least one substantial basis for the estimation of their success. Too much must not be concluded, however, from such comparisons for it must be borne in mind that many factors over which one has little or no control may affect his income. Even the value of the dollar as measured by what it will buy of life's necessities differs materially in different sections of the country. Again a temporary period of prosperity or depression in a given industrial field is no doubt promptly reflected in the salaries paid to employees. In any event it will be necessary to have a much larger number of cases than we are here considering in order to make final deductions. The material, however, is highly suggestive and is believed to be worthy of some consideration.

TABLE 9

Showing the median incomes of the upper, the middle, and the lower thirds of groups Ia and IIa according to personality

PERSONALITY	GROUP Ia	GROUP IIa
Highest third	$3,000	$3,000
Middle third	2,316	2,700
Lowest third	2,076	2,040

TABLE 10

Showing the median incomes of the upper, the middle, and the lower thirds of groups Ia and IIa according to general intelligence

GENERAL INTELLIGENCE	GROUP Ia	GROUP IIa
Highest third	$2,400	$2,856
Middle third	2,580	2,640
Lowest third	2,100	2,856

Of the 21 men in our first group the highest annual income five and one-half years after graduation is $4050, the lowest $1920, and the median income is $2604. Similarly in group IIa the highest annual income four and one-half years after graduation is $7200, the lowest $1740, and the median income $2700. Now dividing the two groups each into three equal divisions on the basis of the ranks of the individuals in personality as shown in table 9 we find the median income of the highest third in group Ia to be $3000, of the middle third $2316, and of the lowest third $2076. Again in group IIa in the same table we see the same tendency for the men with the better personalities to earn the larger incomes. Now turning to table 10 showing the relation between general intelligence and income we find a somewhat different situation. In Group Ia the largest incomes are received not by the most intelligent group, but by the middle third; while in group IIa there is no difference between the median incomes of the highest and lowest thirds.

TABLE 11

Showing the correlations between (1) personality, (2) general intelligence, and (3) school grades and the incomes of the men in groups Ia and IIa five years after graduation

	INCOME AND:		
	Personality	General intelligence	School grades
Group Ia	0.72	0.18	0.32
Group IIa	0.46	−0 18	0.19

The general tendencies of tables 9 and 10 are clearly borne out by Table 11, showing the correlations between income, personality, general intelligence, and school grades.

Since it is entirely obvious that one's personality is in some way closely related to his earning power, we may now ask how far each of the various personal traits is related to one's income. In group Ia where we have the more detailed analysis of personality, we find the 23 traits related to income in the following order and amounts of correspondence:

Accuracy	0.77
Originality	0.75
Address	0.72

Social, civic interest	0.72
Memory	0.69
Enthusiasm	0.68
Motor ability	0.65
Aggressiveness	0.62
Popularity	0.61
Sympathy	0.61
General ability	0.60
Reasoning ability	0.56
Tact	0.56
General information	0.53
Appreciation of humor	0.53
Self-reliance	0.48
Coöperativeness	0.47
Speed in work	0.47
Industry	0.45
Reliability	0.36
Neatness	0.30
Sincerity	0.29
Moral habits	0.23

ORIGINAL WORK

Approximately five years after graduation these two groups of men have produced a total of 22 different items or pieces of original work in the form of 14 scientific articles and technical publications and 8 mechanical inventions or devices that have received more or less recognition. In addition to this 3 men have done some writing of a non-scientific character for the publications of the firms with which they are associated, while each of 2 others has received special honorable recognition in his work. Reference to tables 7 and 8 shows that most of this original work has been done by men above the average in personality in each group although the relation is not strikingly close. The average personality ranks of the 6 men in Group Ia who have made some original contribution is 8 and in group IIa the corresponding average rank is 6. In each group then the average personality of the writers and inventors is slightly above that of the group as a whole. Recalling the traits which enter into the make-up of personality in our experiment (see table 1) we should not expect a closer correspondence than is indicated.

We find about the same tendency to correspondence between original work and intelligence, objectively measured, in group Ia as between personality and original work, but in group IIa the correspondence between intelligence and original production is much closer, the average intelligence rank of the producers being three. In the case of school grades and original contributions the correspondence is closer, the average rank of the contributors in group Ia being 6.3, and in group IIa, 3. This is, no doubt, easily explained by the fact that good grades and original production both depend upon general intelligence plus certain personal traits to guarantee that the mental ability will be made to function.

It has often been observed that scientific investigation, writing, and research are not highly remunerative activities. This observation receives some support from what evidence we have here. In group Ia the average income rank of the writers and inventors is 9.3; while in group IIa it is 7.5. In other words, while these ten men who have already made some scientific contribution in their respective fields are somewhat above the average in their respective groups in general intelligence, personality, and school records, their incomes are not proportionately higher than the average. Here again the joy, satisfaction, and social recognition which the individual gets from such work evidently act as a partial compensation.

THE NORMALITY OF THE TWO GROUPS

It should be stated that the two groups of men who form the basis of this experiment are typical and representative seniors in a typical mid-western engineering school. We have checked their scholarship records very completely and find the average of group I to be less than 1 per cent above the average for all engineering seniors in the university. Group II is about 1 per cent below the average. It is frequently assumed that only the best individuals of a given group respond to such an inquiry as was sent to these men so that the replies are not precisely representative. This is apparently not true of these two groups. The average personality rank of the 21 men in group I who replied to the inquiry is 15 or slightly below the average of the entire group; the average intelligence rank is 12, and the average scholarship rank 14. In group II the case is very similar. Of course, we cannot be sure that those who replied were not the most successful and prosperous men, but the figures just given do not warrant nor suggest such an assumption. On the other hand, they do suggest that the results are true and valid and are representative of the entire group.

CONCLUSIONS

It is possible by means of such a procedure as is outlined in this experiment, with certain refinements, to determine the relationships of the various personal and character qualities which enter into the composition of personality.

It is also possible to determine in a similar fashion the relative *importance of various personal traits for success in different* vocations. There is no reason why the relation of such traits and characteristics as weight, height, hours of sleep, punctuality, eating habits, use of profanity, and scores of others, to vocational choice and success should not be determined within reasonable limits.

A good personality is characterized by such traits as enthusiasm, accuracy in work, aggressiveness, and self-reliance; while such traits as neatness in appearance, appreciation of humor, and moral habits contribute but slightly. It must be borne in mind as has already been suggested that all our conclusions and the preceding one especially, are subject to the *conditions of the experiment*. The group is limited to men of a high grade of intelligence, generally; to men of similar training and experience; and to men largely engaged in fairly similar vocations. A larger number of cases and a longer period of observation of the men in the various vocations may modify somewhat our results.

There is apparently no relation within a group of highly intelligent men, for example, men whose I.Q.'s are above 110, between financial success and general intelligence as measured by an objective test. It is entirely obvious that there is considerable correspondence between general intelligence and financial success if we consider all levels of intelligence from the genius down to the imbecile. It is also possible that within a given restricted field of activity, such as newspaper editorial work, or scientific research, there is some correlation between material remuneration and general mental ability.

In contrast to the lack of correspondence between general intelligence and financial success there is a fairly close relation between personality and income. The relation is indicated by a correlation coefficient of from 0.45 to 0.75.

College grades are somewhat less indicative of the student's ability to earn money than is his personality.

Such traits as accuracy in work, originality, address, social and civic interest, are important factors in determining one's income; while such traits as neatness, sincerity, and moral habits have little to do with the matter.

Such original work as mechanical inventions and scientific publications is largely produced by the most intelligent third of college graduates; the financial reward for such work is not commensurate with the high grade of intelligence necessary for its pursuit.

THE RELIABILITY AND RELATIONSHIPS OF THE COLGATE MENTAL HYGIENE TEST

RALPH K. HOITSMA

Colgate University, Psychological Laboratory, Hamilton, New York

The Colgate Mental Hygiene Test is an extension of Dr. R. S. Woodworth's "Psychoneurotic Inventory" based upon extensive experimentation with normal individuals. The type of response has been designed in the present test so that it is possible to have a quantitative measure of the amount of deviation from normal in each trait. In addition half a hundred items dealing with introversion and extraversion have been added.

The test is in use as a "spotting test" to find those individuals who deviate most from normal in their indirect—and undesirable—emotional outlets. Occupational differences, racial, sex, and age differences are also under study.

The test can be taken in groups as well as individually. Each question (fig. 1) in the large type to the left is answered by making a check mark at the proper place along the dotted line. The check mark may be placed above one of the phrases or between them. The behavior for the past six months only

18. Have you been inclined to work out things alone?	rarely asked for aid	worked mostly unaided	did not hesitate to seek aid	got help in most cases
19. What attention have you given dress and personal appearance?	unconcerned about fashions, etc.		selected mostly for appearance	close attention to dress

FIG. 1. SAMPLE QUESTIONS

There are 75 in B1 and 53 in C1

is considered, although the test may be used to have relatives describe patients as they were in adolescence.

The test is scored by means of celluloid stencils. These show the section of each dotted line in which a check mark indicates that the individual deviates from normal. This section has been determined experimentally from the distribution of check marks made to each question. It indicates that the individual under study is "more so" in the trait than three-fourths of the population studied.

The responses in which the subject deviates are marked with a colored pencil. The total of these colored marks is compared with a percentile chart to show how the individual compares with a normal group. If, for instance, there are 15 colored marks on Schedule C1 he is more introvert than 76 per centum of others. If there are 12 colored marks on B1 it indicates that he is more psychoneurotic than 30 per centum of others.

Complete test forms and percentile charts will be found in the Journal of Abnormal and Social Psychology for July, 1925.

The data to be reported were compiled and checked by the following students: F. S. Bambace, V. N. Barrington, J. W. Baumann, H. P. Bromfield, H. D. Bryant, D. DeNoyelles, I. H. Faulkner, L. B. Goodenough, A. W. Graning, A. R. Hedeman, H. K. Hunsicker, C. W. Jennings, S. B. Jones, J. S. Kaye, M. D. Livermore, W. C. McConnell, P. R. Martino, H. C. Millard, R. W. Moncrief, S. S. Munro, H. H. O'Connor, W. R. Patterson, D. Posson, W. E. Rice, M. L. Rofrano, H. I. Roll, C. J. Rutishauser, G. F. Swartz, H. Seltenrich, J. K. Smith, W. Spaid, L. C. Stark, F. Stanton, J. C. Stuart, M. B. Wilcox, Paul Fry, and Miss Muriel Cutten.

METHOD

This study of the reliability of Schedule B1 and C1 was carried on to find the possible error on the part of subjects in rating themselves and the general reliability of the schedules as tests of emotional stability. Several methods were used to test the reliability of the schedules:

1. On Schedule C1 there are five sets of duplicate questions. These were compared with one another.

2. In Schedule C1 the first half was correlated with the second half. In Schedule B1 the scores in the first half of each section were correlated with the number of symptoms on the second half of each section for the entire schedule. The two schedules were also correlated with one another and with scholastic attainment.

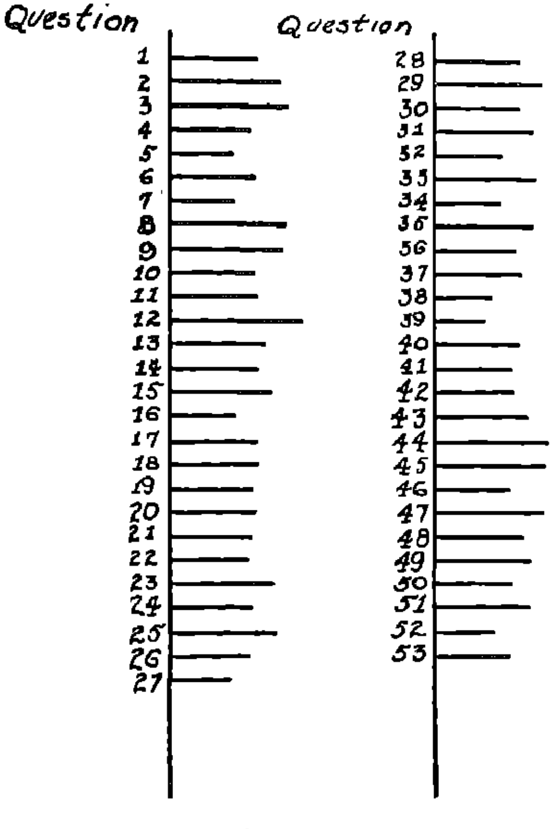


FIG. 2

3. In addition to this a group of 88 subjects was selected at random to take the inventories a second time two weeks after they were first tested.

DATA FOR SCHEDULE C1 (INTROVERSION-EXTRAVERSION)

The correlation between the total number of symptoms the first and second time of taking the inventories was +0.674 ± 0.039. This correlation shows a fair degree of reliability. (All correlations are computed by Pearson's product-moment method.)

The first and second half of the schedule correlated +0.450 ± 0.003 with each other. This correlation is not as high as the one found when the first and second times of taking were correlated but shows a definite relationship between the two parts, although a higher one is to be desired.

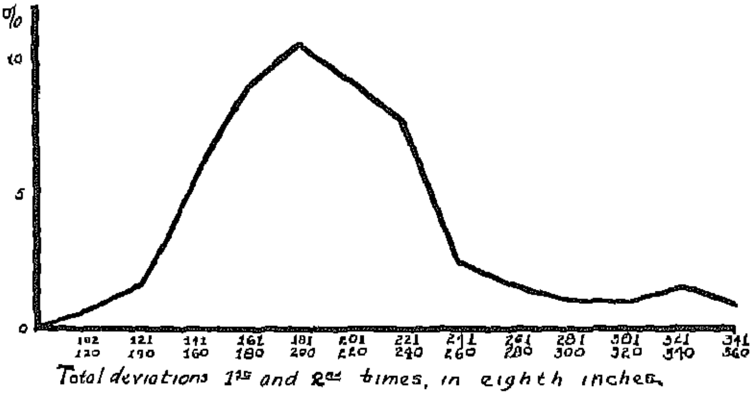
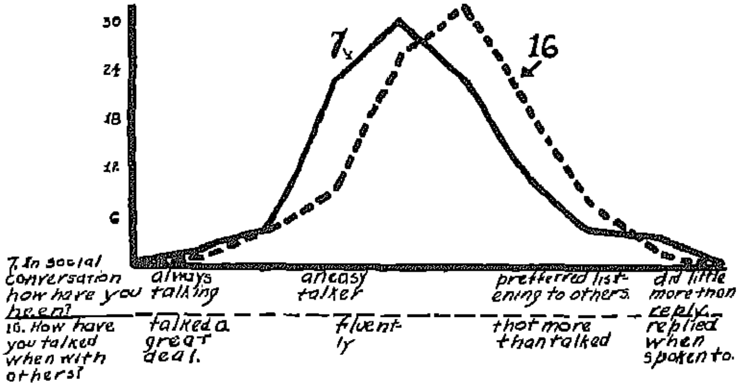


FIG. 3

A study of the inventories taken two weeks apart shows the average misplacement of the check marks for each question to be as shown in figure 2. The length of each line represents the actual average deviations of the second from the first check mark.

Figure 3 shows the distribution of the total of these deviations appearing on each subject's schedule. All deviations on each individual's schedule were totalled giving the total number of $\frac{1}{8}$ inch deviation which appeared on the schedule. This curve shows an approximately normal distribution of total deviations made by those who took the inventory the second time. Some individuals are much more reliable than others, the best deviating on the average only $\frac{1}{8}$ inch and the worst $\frac{3}{4}$ inch.

Figures 4a, 4b, 5a, and 5b show a comparison of the duplicate questions on Schedule C1; 4a, 4b, and 5a are for 488 men subjects and 5b is for 258 women subjects. There are five sets of duplicate questions. The wording of each set of these questions is very similar. In some sets there is a difference in the number of qualifying phrases beneath the rating lines, and in all the "good" ends are reversed.



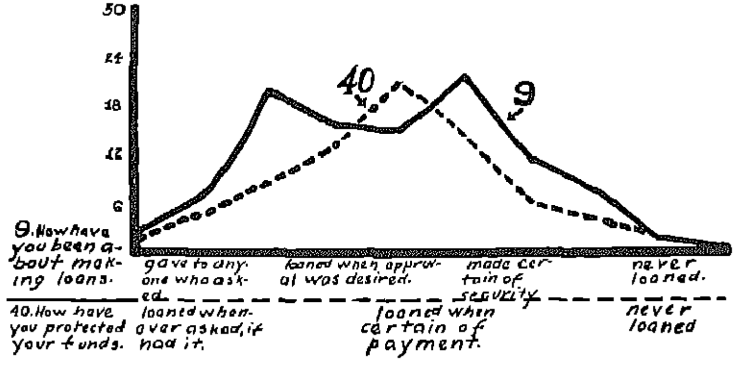
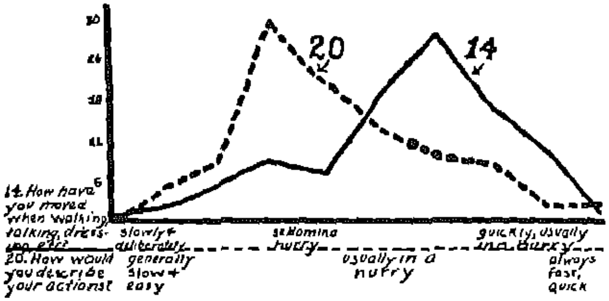


FIG. 4a (Upper)
FIG. 4b (Lower)

The questions compare favorably. If the responses were haphazard the distribution curves illustrated would have approximated normal distribution curves more closely than they do. The fact that they tend to follow the phrases rather than a normal distribution is a point in favor of the method.



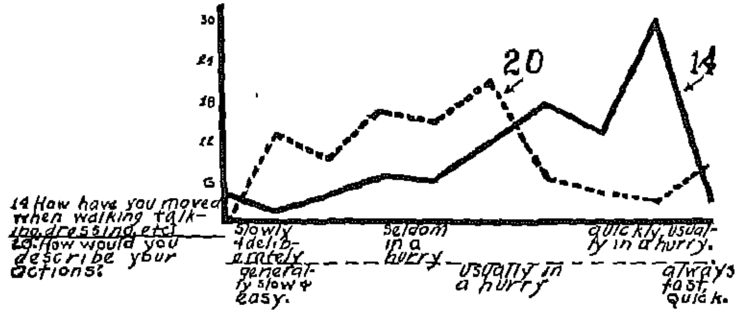


FIG. 5a (Upper)
FIG. 5b (Lower)

Further, if the responses had been largely haphazard the distributions for each question would have been practically the same. To illustrate this point figure 5b may be consulted. In both these questions the high point of the distribution is over the phrase "usually in a hurry" rather than being in the middle of the line for both questions as one would expect if the curve approximated a normal distribution curve. In figure 4b for questions 9 and 40 there is a considerable difference in the curves. There are two high points on the curve for question 9 and one on the curve for question 40. On the questions themselves the wording is very similar but on the rating line question 9 has four qualifying phrases below the line while question 40 has three. This goes further to show the relationship between the distribution of responses and position and number of qualifying phrases below the rating line.

Figure 5b shows similar curves for questions 14 and 20 with the high points of the curves again following the qualifying phrases. Both these questions have three qualifying phrases beneath the rating line. Question 14, however, measures 1½ inches from the middle of the phrase, denoting a symptom of introversion to the middle of the middle qualifying phrase. Question 20 measures 2½ inches for the same distance. This shows a difference of 1 inch. These two middle qualifying phrases show somewhat opposite traits. In question 14 the middle phrase is "seldom in a hurry," in question 20 the middle phrase is "usually in a hurry." The high points of the two curves are 1½ inches apart. This indicates a relation between the distribution of responses and the qualifying phrases on the rating lines. The distance from the introvert end to the high point of the distribution of the responses of question 14 is to the corresponding distance for question 20 as the distance of the middle phrase from the introvert end of question 14 is to the corresponding distance for question 20.

Questions 7 and 16 in figure 4a show very little difference. The same general conclusions as in the above paragraphs hold true in these questions. The high points on both distributions are over approximately the same point in regard to the meaning of the qualifying phrases.

These data show that Schedule C1 is fairly reliable. The correlation of +0.45 between the first half and second half of the schedule show that an abbreviated schedule cannot be used and the r of +0.69 after two weeks shows some revision is needed in this schedule. A valuable hint in revision can be gained from the material on duplicate questions. These questions show that great care used in selecting the qualifying phrases beneath the rating lines would greatly increase the reliability of the test. The data from the second testing can be used to find out which phrases are the unreliable ones.

No data have been compiled as to the reliability of the test with women subjects. If, however, the differences which show between the two sexes in the comparisons that were made on duplicated questioning hold true in all comparisons a difference is likely to appear here also. The women subjects have shown a greater tendency toward introversion. This may be because they have rated themselves more closely or perhaps they are more honest with themselves in their rating. Data are being gathered on this point by Miss Muriel Cutten and will be published later.

DATA FOR SCHEDULE B1 (PSYCHONEUROID)

The average deviation for each question on B1 is graphed in figure 6 as for C1 in figure 2. An average of all these questions gives 3.8 eighth inches, or very nearly $\frac{1}{2}$ inch, as the average distance which one may expect the check mark on the second test to deviate from the mark made on the first test.

The disparity in questions 100 to 132 is the greatest in the schedule. In the case of questions 401 to 407 there is little variation. This seems to be accounted for in the fact that there was little evidence of a normal distribution on these questions which centered around symptoms of hysteria. The check marks were bunched at one end of the rating line. An inspection of this figure shows striking differences among the questions on B1 as to the accuracy of the subjects in rating themselves on successive tests.

A correlation between the total symptoms on B1 the first time of taking and the second time showed r $+0.85 \pm 0.019$. Another correlation was made leaving out section IV of the schedule. This showed an r of $+0.815 \pm 0.023$. It is evident that the lack of disparity in section IV brings the reliability of the whole schedule to a higher point when it is taken as a whole. The first half of Schedule B1 correlated $+0.79 \pm 0.02$ with the

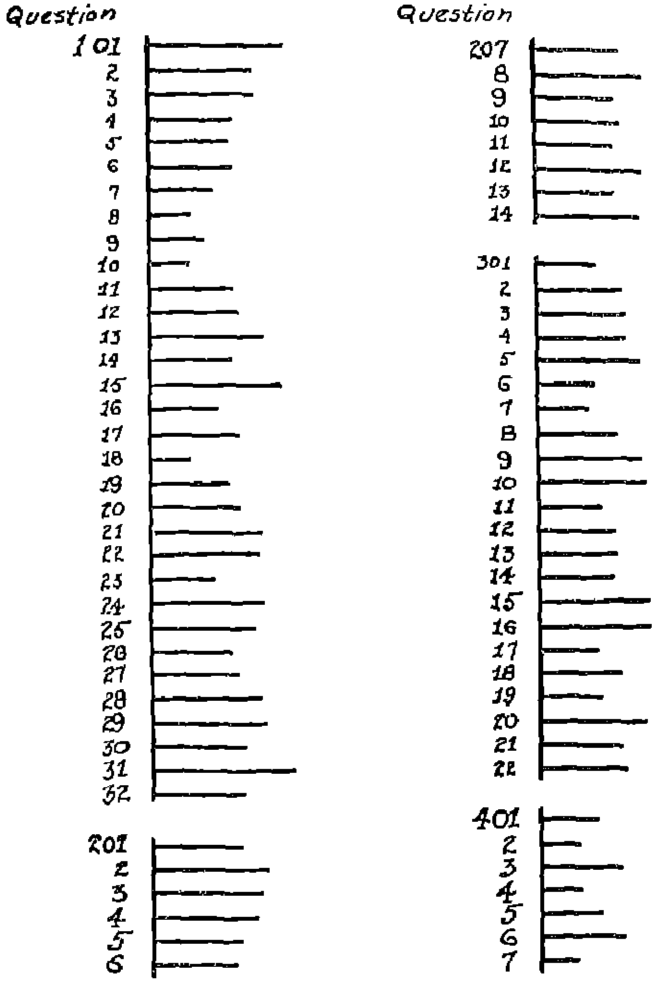


FIG. 6

second half of the schedule. The scores on section I of this schedule (Psychasthenoid) were correlated for the two times of taking the test. *r* in this case was equal to +0.746 ± 0.03.

MISCELLANEOUS RELATIONS

Some interesting facts were found in correlating the schedules with scholarship, intelligence, and with one another.

The scores made on the Thorndike Intelligence Test were correlated with the total number of symptoms of introversion and were found to have no relationship. Two hundred and eighteen subjects were used in making this correlation; r was found to be equal to $+0.02 \pm 0.009$. It is evident that a high score on the Thorndike Test does not mean that the subject will have a large number of introvert symptoms. In this same connection, however a correlation between the number of symptoms of introversion and scholarship gave r equal to $+0.35 \pm 0.069$. A student having a larger number of symptoms of introversion than the average student is also likely to rate higher in scholastic attainment. It is possible that Schedule C can be combined with an Intelligence test and weighted so that scholarship can be predicted much more accurately than with either alone. Experimentation on this is under way.

By using one end of the rating line of C1 as a symptom of introversion and the other as a symptom of extraversion, the number of symptoms of introversion and extraversion were obtained. These symptoms were correlated with one another.

The correlation between introvert and extravert symptoms in the group of subjects who took the test twice was -0.36 ± 0.06. In the case of 288 men subjects r was -0.22 ± 0.141. This relationship was somewhat higher with 268 women, r in this case being -0.45 ± 0.003. The number of extraversion symptoms correlated -0.19 ± 0.078 with scholarship in the case of men.

No relationship was found between the total number of symptoms on B1 and scores on the Thorndike Intelligence test. Records of 203 men subjects were used in making this correlation; r equalled $+0.008 \pm 0.047$. With women this r was -0.12 ± 0.06.

The correlation between total symptoms on Schedule B1 and scholarship was $+0.074 \pm 0.082$ with men. The same ecords were used in making this correlation as was used in the correlation between scholarship and introversion. r was also computed between the number of neurasthenoid symptoms and scholarship and the number of psychasthenoid symptoms and scholarship. In the former r was $+0.085 \pm 0.08$, in the latter $+0.019 \pm 0.08$. These are both the equivalent of no correlation.

A correlation between the total number of symptoms on Schedule B1 and on Schedule C1 showed a relationship of $+0.49 \pm 0.02$. A student high in the number of symptoms of introversion also tends to run high in psychoneurotic symptoms.

These correlations seem to show that the schedules can be made into a valuable aid to administration problems and the selection of personnel. Findings on this will be given in later papers from the Colgate University Psychological Laboratory.

CONCLUSIONS

These data show that the schedules are reliable enough to use in making group comparisons, such as comparing one class with another or one age with another or even smaller divisions. When one attempts to draw conclusions, however, or to establish a difference between individuals when one has 15 symptoms of introversion and another has 17 symptoms, obstacles are certain to be met.

There is evidence in the data presented in this paper that some of the questions in both schedules are highly unreliable and need revision. Section IV of Schedule B1, the section dealing with hysteroid symptoms, is of little value in its present form. The lack of hysteroid symptoms on the part of subjects may be data showing that college groups are composed of people tending more toward introversion than the average person and as a result do not develop the neurosis of extraversion or hysteria.

From the data at hand Schedule B1 is shown to be more reliable than Schedule C1. The data indicate a relationship between the number of questions in the schedule and the reliability. Schedule C1 is two-thirds as long as Schedule B1 and not quite two-thirds as reliable as Schedule B1.

TYPES OF WORK

A Note on Vocational Interest

DOUGLAS FRYER

Assistant Professor of Psychology, New York University

In vocational guidance and personnel selection practical devices for personality and occupational analysis are frequently used, of necessity, in the absence of scientifically established methods of work. A concept of practical utility, often given considerable prominence in vocational prognosis, is embodied in the words "types of work." This paper deals with types of work in the study of the worker's interests. The question is raised: How valuable is this concept of types of work? Is it of prognostic significance for vocational guidance or personnel selection?

In 1920 the writer determined to put this concept of types of work to the practical test of the vocational office in the study of interest. The purpose was to establish the prognostic value of interest in types of work.

Occupations were classified into two large groups as they seemed to appeal to interest: (*A*) the *humanics*, work with people and (*B*) the *mechanics*, work with things. Later the humanics was subdivided into two groups: *concrete humanics*, work directly with people, and *abstract humanics*, work indirectly with people; the mechanics was subdivided into two groups, *concrete mechanics*, work directly with things and *abstract mechanics*, work indirectly with things, or with the symbols of things. This classification appears below as it was presented verbally by the vocational counselors to a group of 200 counsel subjects and as presented in questionnaire form to 513 employment applicants.

Classification of types of work

A. *Humanics*, work with people:
   1. *Concrete humanics*, work directly with people, influencing, advising, directing people, like the buyer, the salesman, the executive in direct contact with the employed staff, etc.
   2. *Abstract humanics*, work indirectly with people, promoting and forming policies, ideals, and organizations for people to work or live by, like the publicity man, the advertising writer, the journalist, and the executive not in direct contact with the employed staff.

B. *Mechanics*, work with things:
   3. *Concrete mechanics*, contact directly with the objects of work, manipulating the tools and machinery, dealing with the things themselves, like the mechanic, the construction worker, the farmer and a great deal of the work of the engineer.
   4. *Abstract mechanics*, work with language and figures, dealing with the symbols of things, like the accountant or bookkeeper, the draftsman and the bank teller.

This classification of types of work was tried out as the regular part of the counsel process in a vocational office in New York City. According to the judgments of the subjects, following a brief description of the types of work, these 200 subjects divided themselves on the basis of their interests as shown in table 1.

Quantitatively this table appears to indicate that individuals can usually make generalized interest distinctions of the sort called for in classifications of types of work.

A qualitative analysis of this experiment based upon the reports of the counselors gives another aspect. It was evident in the discussions that there was a confusion on the part of the subject whether he was estimating his interest in the type of work or his ability in that type of work. Confusion between interest and ability was common. It became evident to the counselors that it was practically impossible for the subject to distinguish in his estimate between ability and interest with any high degree of reliability.

Furthermore, the subject often failed to get the meaning of the general concept represented in the particular type of work. If he had had experience in *any* of the occupations included in the type under discussion his judgment seemed to be based upon that. His judgment seemed, in so far as it pertained to interest, to be based upon the feeling accompaniment to the thoughts of this particular task. This is in accord with our understanding of conscious processes. There can be very little feeling accompaniment to the concept. But when the concept is represented in consciousness by the idea or percept there can be any amount of affective elements.

TABLE 1

GROUPS	NUMBER		PER CENT
Humanics:		111	55
Concrete humanics	57		
Abstract humanics	23		
No distinction*	31		
Mechanics:			
Concrete mechanics	14	72	36
Abstract mechanics	47		
No distinction*	11		
Both †		17	9
Total		200	100

* No distinction between the concrete and abstract was requested.

† Seven of these individuals designated their interest as in both abstract fields, three designated it as in both concrete groups, five as in both abstract mechanics and concrete humanics, while two were unable to chose.

Lack of knowledge of the occupations frequently caused the subject to come to exact judgments, of precise but highly superficial nature. Such statements as these would follow the counselor's description of types of work: "Oh I like to deal with people. . . . . I like to have the finished product in my hand. . . . . I want to be out in the open in mechanical work. . . . . I have a mechanical bent.

. . . ." But below this superficial judgment there would be absolute lack of understanding of the work involved. A discussion of the nature of work would bring an "informed" ignorance that would make any judgments as to types of work impossible. Fortunately for the quantitative record given above the subjects' judgments were recorded in the first part of the interview or it would have been impossible in many cases to secure any record.

The quantitative results appear lacking in significance when subjected to a searching analysis. These subjects do not seem to be able to generalize upon their interest in types of work, except in the most superficial fashion.

There is offered, however, a piece of suggestive evidence. The classification of types of work offered the subject a psychological picture of the vast scope of occupational endeavor. It became the starting point for discussion of specific occupations and it served as the basis for later occupational orientation. While it lost its meaning as types of work in the discussion it was suggestive to the subject in commencing a vocational analysis from the point of view of interest.

Further study of the significance for vocational prognosis of the concept of types of work was later carried on with a group of 513 subjects representing over 100 occupations, all applicants at a commercial employment agency in New York City. This study was made by means of a questionnaire containing questions relating to interest in the types of work classified above. Other questions were asked upon the questionnaire and of the 513 returns only 440 answered the questions relative to interest in types of work, indicating probably that the distinctions called for were too difficult for some. Of the 440 subjects for which answers were received 209 stated that they preferred work in both the humanics group and the mechanics. Thus of the 513 subjects 53 per cent in all failed to make a distinction between these two large groupings of types of work.

Of the 440 subjects answering the questions upon types of work the results are given in table 2.

Forty-eight per cent of the group that responded to the questions failed to distinguish between their interest in work with people or with things. Two sets of questionnaires were used with this group, the first set failing to contain the abstract and concrete subdivision of the humanics, thus accounting for the large group making no distinction there.

The quantitative results from the two samples, the counsel subjects and the employment applicants, are distinctly different. Only 9 per cent of the former sample failed to make the distinction between the humanics and mechanics groups while fully 50 per cent failed to do so among the employment applicants. There are two possible explanations of this difference. With the counsel group a discussion of the meaning of the distinctions, with specific occupational illustrations, was possible. The subjects could thus be influenced to some sort of choice and the interpretation of the counselor could play a large part in recording the decision of the subject. The qualitative analysis of the counsel subjects suggest that the explanation has considerable factual basis.

TABLE 2

GROUPS	NUMBER		PER CENT
Humanics:		146	33
Concrete humanics	63		
Abstract humanics	42		
No distinction*	77		
Mechanics:		85	19
Concrete mechanics	43		
Abstract mechanics	35		
No distinction	7		
Both humanics and mechanics		209	48
Total		440	100

* No distinction called for in most cases.

The other explanation is that better educated and more intelligent subjects are able to make these general distinctions called for by the classification of types of work. The sample composed of counsel subjects represents a distribution of intelligence on the Army classification of three per cent in the inferior intelligence groups, seventy per cent in the superior intelligence groups and the remainder in the average group, while for the employment group the intelligence distribution is much lower, with the median just over the line of the high average intelligence group. The latter group, composed of individuals drawn from over one hundred occupations appears to be more representative of the general male population in intelligence. A similar difference, but not so great, appears between the two groups in education.

In the sample composed of employment subjects, where there are sufficient numbers for a comparison between those making a choice for humanics or mechanics and those choosing both, there is a slight indication that the more intelligent and the better educated are inclined to choose *both* work in the humanics and the mechanics. Apparently education and intelligence are factors influencing the decision by causing a disinclination to choose between the two types of work. It is possible that greater knowledge of possibilities in both fields enters to prevent the choice among the better educated group.

Out of the intelligence and educational analysis of the quantitative results of both samples there appears a distinction between the abstract and concrete work. Intelligence and education are factors in causing the choice of abstract rather than concrete work. The more intelligent and better educated subjects making the distinctions called for choose work of an abstract nature to a marked degree. This is found to be true in both the humanics and mechanics group.

SUMMARY

From this attempt to grapple with the problem of interest in types of work the following conclusions arise:

1. Individuals are unable to make any generalized voca-

tional interest distinctions of significant value for vocational prognosis. Neither intelligence nor education are important factors in such decisions. Interest, which may in this instance be defined as the feeling of an individual toward certain occupational endeavors, cannot be expressed toward anything so general as a type of work.

2. Types of work based upon fields of occupational interest, used solely as a logical classification of work, can well serve as a means of orientation in vocational counsel where the information process must necessarily be from the general to the specific. On the other hand, as psychological types of work, these classifications appear to be of as little significance as the classifications of types of mind, a concept which has never yet been able scientifically to establish its right to exist.

PHENOMENAL MEMORY IN ITS BEARING UPON VARIOUS MENTAL TESTS

A CASE REPORT

FRANCES E. OTIS

Letchworth Village, Thiells, New York

Among the patients at Letchworth Village at the present time is one who merits particular study and who well repays an individual for the time spent in working with him. Mark, a nineteen-year-old boy, has a remarkable memory and recalls with facility anything which has been taught him—dates, names, places or events. Geographical names of obscure and little known places, dates of significant or insignificant historical facts, names of important or unimportant persons, if once known are never forgotten.

The question immediately arises as to the nature and cause of the mental defect or variation in such a boy. At once we arrive at a blank wall. There seems to be slight indication of the underlying cause in this case and a question exists as to the nature of the disorder. The family are of German Jewish origin of fairly good social class. There seems to be nervous instability on both paternal and maternal sides, the father having had an involutional melancholia and the mother having been neurotic. Nothing unusual is known of the prenatal history of Mark himself. Early developmental history shows definite retardation—the child not walking until one and one-half years or talking until three and one-half years old. The probabilities are that the defect, if defect it really be, is not a secondary one. But is it a true mental defect, or is it a psychosis, or perhaps a brain disease of obscure origin? The diagnosis in the case is extremely doubtful, but aside from that, the fact of his unusual talent and his social inadequacy still remain.

THE JOURNAL OF APPLIED PSYCHOLOGY, VOL. IX, NO. 3

The patient was a behavior problem from an early age—was careless about his personal habits, untidy, and difficult to work with. He was sent to kindergarten at six and to public school at seven. Upon the death of his mother when he was ten years of age, he was sent to a boarding school since none of his relatives would care for him. Here he attended public school for over five years. He was finally sent to a Farm School in New Jersey where he stayed but a month. Supervision in this school being more lax than at boarding school, Mark immediately fell into difficulties trying to make decisions for himself. At the boarding school he showed anti-social sex tendencies, was easily led, and demonstrated his inadequacy to decide for himself or carry out plans. After his removal from the Farm School, his uncle procured work for him in a paper concern, but this experiment proved a total failure. He was then sent to Bellevue for observation and after four weeks was committed to Letchworth Village.

Our next move is to discover what has been observed in this boy since his admission. His physical examination upon admission was essentially negative. He has the stooped posture and somewhat shuffling gait characteristic of a mental defective. His personality charting shows him to be of a more or less cheerful and calm temperament, sluggish almost to the point of laziness. He is alert to his environment, aware of what goes on around him, is fairly sociable, coöperates almost too well, and is respectful and tractable. He is, however, a bit egotistic, and is underhanded in many of his dealings. He is, moreover, a definite homosexual. This is quite evident both from his previous history (the anti-social sex tendencies referred to being an attack upon a younger child) and from his institutional life. During his stay at Letchworth Village he has been involved in several affairs of a grossly homosexual nature. Aside from these definite manifestations, he shows the general characteristics of a homosexual. He is subservant, overly anxious to please and apt to do more than he has been asked to do. He fails utterly in tasks which are of fairly long duration, though of a simple nature, which carry him away from those in authority but succeeds in those tasks which are quickly done and involve a personal contact. For this reason he is of little real value in the institution though he works fairly well around the cottage, and aside from a few instances where his anti-social sex tendencies have been displayed, causes little real difficulty. He has a wide general information, reads the daily newspapers intelligently, can converse in a manner which far surpasses that of the ordinary defectives and shows a relatively good command of language.

With a boy possessing a capability such as this, what can be expected from the use of various mental tests? Would this extraordinary capacity affect the tests and to what extent? If our results show that this ability tends to give a high mental age do we then agree that he is not in any sense a mental defective despite his apparent failure along other lines?

Upon admission Stanford-Binet test gave Mark a mental age of 12^{11}, Intelligence Quotient 81. This result, however, can be immediately discarded. His uncle not wishing him to be rated as a mental defective coached him on the answers to the Terman test and due to his remarkable memory Mark was able to secure this rating. Here is a practical demonstration of his unusual ability, but no rating of his true mental capacity. It is interesting to note that on a mental test given February 29, 1920, two years prior to the one above mentioned his mental age was the same 12^{11} and his Intelligence Quotient 90 due to lower chronological age. A mental test given October 10, 1921, shows mental age 13^{0}, Intelligence Quotient 81. The mental age then appears to be constant, showing that there has been no mental growth while Intelligence Quotient has steadily lowered due to a higher chronological age.

From such an outline of history and behavior what then is the progress for results on mental tests aside from the Stanford-Binet? Using as representative tests the Porteus, Pintner-Paterson, Heally II, Trabue Completion tests, let us make a tentative prognosis bearing in mind the behavior noted and the unusual memory displayed. Such a tentative prognosis would indicate unquestionably a low score on the Porteus test. Such lack of planning capacity and ability to meet and cope with situations could bring forth only a low score here. The Pintner-Paterson Performance median would also undoubtedly be low. His failure at employment in the paper concern (it being impossible to teach him to tie a package properly) would give evidence of this. On a few of the tests in this scale, however, he promises to give good scores. He should show special ability on the Knox cube test if general memory is any indication. The fact that he is alert and well oriented socially would also indicate a good score on Picture Completion. His excellent immediate and retentive memory should bring forth, too, a high score on the substitution test—that being essentially a test of easy learning and memory.

The prognosis for Healy II is extremely good. Mark's excellent social orientation and wide general knowledge gained from reading and fair education in addition to the fact that his early environment was New York City all indicate good prospects in this test.

The Completion tests present more or less of a problem in prognosis. The lack of judgment apparent in the patient should lower the result. On the other hand, his language proficiency and wide enough knowledge to grasp the general idea seem to indicate good results. Balancing the two, it would appear that the results on Completion tests should be fair if not good.

The actual results on the tests correlate fairly well with the prognosis. The performance median is 7 years, 8 months, quite as was expected. The score on Mare and Foal test was 5 years on errors, 7 years on time. Goddard Form Board gave mental age of 8 years. He was unable to complete Two Figure and Casuist boards, having no sense of form discrimination. He reversed the arm on Manikin test, a result which is a bit surprising in view of the fact that a good social orientation was assumed. He could not complete Feature Profile test. Scores on Ship Test and Picture Completion were considerably higher, 9.5 and 11 years respectively. Misplaced two upper adjacent pieces in Ship test and placed one absurdity in Picture Completion. As was expected he made his highest score on Knox cube test, mental age 12. A failure of attention rather than memory may be the reason for the score on this test being no higher. The assumption might have been made that he would have had a perfect score with such unusual memory, and with perfect attention such might easily have been the case. He failed the second movement in Adaptation Board, giving him a mental age of 7 years. The substitution test showed mental age of 8 years with three errors. On Healy A, 88 moves and 4½ seconds were required for completion, giving him a mental age of seven years.

The actual mental age on Porteus test was 5^0. He completed the five year test with no difficulty, failed the sixth year entirely and succeeded with seventh year on second trial. This result coincides with prognosis—emotional instability, poor judgment and lack of planning capacity are all shown up in results.

A mental age of 9^0 was the result of his work on the Healy II, showing fairly good social adaptation, quite in accordance with prognosis. In working he placed two absurdities as opposed to five logical placements. He was not particularly open to suggestion for he retained these logical placements even when questioned concerning them and gave logical reasons for each placement.

On Completion test Alpha and Beta Mark gained a mental age of 11^8 and 11 respectively with Intelligence Quotients of 73 and 68. Here we see a decided apparent inconsistency with results of accompanying tests. On tests which are supposedly tests of logical reasoning this boy who shows little actual ability along this line achieves a higher score than in any other tests. This inconsistency may however be alleviated in some measure when we take into consideration the fact that the completion test include, as we observed before, language ability and are objective rather than subjective. That is to say, the patient can reason fairly well on things outside of himself and outside of actual situations, but when brought face to face with realities fails.

The next test attempted by the patient was a so-called Qualitative test, unstandardized, wherein an attempt was made to test separately the qualities of attention, memory, comprehension, judgment and retentive memory. The maximum score in each dimension was 12, the idea being to compare in a group of patients one patient with another in respect to the various abilities. On this test Mark scored as follow Attention 9⅓, Memory 10½, Comprehension 11½, Judgment 8¾, Retentive Memory 12. The poorest results then are on judgment and attention, highest on retentive memory. This maximum score on retentive memory has, however, been attained by many others of approximately the same or lower mental age without any special ability, so it shows no particular indication in this case. The failure on the tests of memory was in the auditory digit span. He repeated seven digits forward which places him in the fourteen year level according to Terman. He failed 8 digits forward. His comprehension was good, misconstruing as he did but one direction. We can not, nevertheless, immediately say that Mark's comprehension is as good or better than his memory. Though his score in this test is as high on comprehension as on memory, yet there have been no mental age standards set on these tests and the probable deduction is that if this were done, the mental age on the comprehension group would be lower than on the memory group. A low score on memory may indicate then as high if not higher mental age than a high score on comprehension. This seems consistent in view of the fact that the mental age on these tests of auditory memory goes as high as fourteen or sixteen years.

On the Pressey Primer scale which forms part of the tests on judgment Mark scores a mental age of 7^0. On Dot Pattern test he had 11 correct markings out of possible 25, completing the entire 25 in the given time. On Classification test he succeeded in 16 out of possible 25; on Form Board test passed 15 out of 25 and in Absurdities 15 out of possible 25 again. In each case he completed the entire 25 in the given time. He showed extreme scattering on failures in Dot Pattern test, failed 2nd pattern and passed 18th and 19th. He showed, too, much variation on Classification test. Form Board failures were very consistent in that nearly all occurred as test grew progressively more difficult. In Absurdities the same consistency is shown. The mental ages on the four tests are 6^4, 8^0, 7^6, 8 respectively, composite mental age 7^9.

A question may arise then as to whether or not there is a discrepancy between the mental ages attained on the completion Alpha and Beta and that attained on the Primer Scale. Broadly speaking, all of these are tests of judgment, but they vary so much in kind that no actual discrepancy can be validly asserted. Tests I and III of the Primer scale (Dot Pattern and Form Board) are tests of Form Discrimination in a large measure, while tests II and IV (Comparison and Absurdities) are tests of Logical Relations. At the same time, the factor of memory and easy learning enters into the Form Board test and the factor of social orientation enters into the test of absurdities and each exerts a notable influence.

The composite mental age attained by Mark on the entire series of tests given (omitting the qualitative necessarily) is computed as follows:

WEIGHTING	TEST	M A.	W + g × M.A.
2	Performance	7–8	15–4
1	Beta	11	11
1	Alpha	11–8	11–8
1	Porteus	5-6	5–6
1	Healy II	9–9	9–9

Composite mental age 8 years, 5 months.

This composite rating is certainly much more satisfactory than the padded mental age of 12^{11} procured on Terman.

The mental tests then do not startlingly demonstrate any unusual capacity along one line, but do unquestionably show Mark's general social inadequacy. That is to say, the unusual memory present in this patient does not in itself alter the scores on the mental tests to an appreciable degree.

From the results of these tests and from behavior noted, it is quite apparent that Mark will always be a custodial case. Fairly capable under direction, coöperative and tractable, easily adjusted as he is, he is of little care in the institution. Out of it, he would be totally at a loss, unable to cope with the slightest difficulty and would undoubtedly manifest to a decided degree his anti-social tendencies. Undecided then as the diagnosis is in the case, the fact remains that Mark possesses remarkable ability along one line, coupled with general unfitness to conduct himself efficiently in extra-institutional life.

NOTES AND NEWS

A. J. Snow, Consulting Psychologist to the Yellow Cab Company, Chicago, Ill., has recently been retained in the same capacity by Kronn and Dato, the largest real estate dealers in Chicago. His new duties consist of the organization of the personnel department and the devising a test for the selection of several thousand salemen. In addition to this the Yellow Cab Company has asked him to reorganize the training school for drivers and to organize a sales school with an annual budget of sixty-five thousand dollars.

Dr. Helen P. Woolley, Psychologist of the Merrill-Palmer School, Detroit, has been appointed Director of the Institute of Child Welfare Research, Teachers College, Columbia University, New York City. She takes the place of Dr. Otis W. Caldwell, Principal of the Lincoln School. In addition to this appointment Dr. Woolley has been appointed Professor of Education in the Faculty of Education.

Dr. Bess Cunningham has been elected Supervisor of the Educational Clinic and Assistant Professor of Education, Teachers College, Columbia University. Dr. Cunningham has served as Secretary of the Institute of Child Welfare Research, Teachers College.

Under the auspices of the Child Study Association of America, Inc., a conference will be held at the Hotel Waldorf, New York City, from October 26 to 28, in which will be discussed many topics of vital interest to parents and all who are interested in the problems of the family as it exists today. Among the speakers will be Dr. Bernard Glueck, psychiatrist; Dean Russell, Teachers College, New York City; Helen T. Woolley, the new Director of the Institute of Child Welfare Research, Teachers College; Miriam Van Waters, author of "Youth in Conflict"; and Dr. Beatrice Hinkle, author of "The Re-creating of the Individual." The public is invited to attend this conference.

BOOK REVIEWS

CHARLES W. ODELL. *An Annotated Bibliography Dealing with the Classification and Instruction of Pupils to Provide for Individual Differences.* Bulletin No. 16, 1923. Bureau of Educational Research, University of Illinois, Urbana. Pp. 50.

In this monograph we have three hundred forty-six titles listed. This list indicates something of the present interest of educators in the study of individual differences. But since the appearing of Odell's bibliography in November, 1923, two additional ones in the same general field have been printed. The first of these appeared in the Twenty-Third Yearbook of National Society for the Study of Education and the second in the Twenty-Fourth Yearbook of the same Society. The combined bibliographies record approximately thirteen hundred titles. There are relatively few duplications and when taken collectively they supply the reader with a fairly complete record of material relating to the topic of individual differences.

These bibliographies have the merit of being carefully annotated, thus greatly increasing their value. It would be a distinct contribution to this general field of education to have the three bibliographies combined and carefully indexed.

WALTER S. MONROE. *A Survey of the City Schools of Marion, Illinois.* Bulletin No. 21, 1924. Bureau of Educational Research, University of Illinois, Urbana. Pp. 60.

In this monograph we have the report of a survey of the school system of a small city. The findings of the survey staff are presented under the usual captions, including the city's investment in elementary education, the condition of buildings and equipment, the teaching staff and quality of instruction, the course of study and achievements of school children. In a final chapter certain specific recommendations are presented.

The report has the merit of being brief, setting forth only such facts as are beneficial to the community concerned. The recommendations are made with the thought of securing the maximum of improvement in the local system with the minimum of expenditure.

WALTER S. MONROE. *The Constant and Variable Errors of Educational Measurements.* Bulletin No. 15, 1923. Bureau of Educational Research, University of Illinois, Urbana. Pp. 30.

In this brief monograph the author discusses constant and variable errors in educational measurements, and points out their influence on such derived measures as the average, median, standard deviation, and coefficient of correlation. In the course of the discussion there is presented evidence of some five different sources of constant errors in educational measurements. These sources of errors may be summarized as follows: (1) The successive administrations of the same test or its equivalent form; (*a*) the lack of equivalence of duplicate forms of a test; (3) instruction functioning as coaching; (4) the subjective factor in the scoring of the tests; (5) the attitude of the pupils toward the tests. The author points out the possibility of constant error resulting from each of these sources and then adds that it is difficult to determine the exact magnitude of these errors. It is only possible to state that the error probably does not exceed a certain amount for we cannot be sure that only one source of error is operating in any single trial of a test.

A careful examination of the scores of any test will reveal evidence of the presence of variable errors of measurement. This becomes quite apparent when the same test is repeated after a short interval and the two scores of the same individual pupils are compared.

What should be the effect of a knowledge of the existence of such errors upon the movement for scientific measurement in education? It is not the author's purpose to advise the dicontinuance of the use of educational tests. It is rather to point out to those using such tests the limitations of the measures yielded by their use. It is certainly worth while for both the administrator and the school statistician to recognize the possible presence of one or both types of errors in every application of the standard tests.

WALTER MONROE AND DORA KEEN MOHLMAN. *Training in the Technique of Study.* Bulletin No. 20, 1924. Bureau of Educational Research, University of Illinois, Urbana. Pp. 68.

In recent years our attention has been directed to the conviction that study habits form an important educational objective. In this monograph we have presented a method of procedure looking toward training the child in good habits of study. After selecting a group of high school students who were doing poor work a careful study was made of the causes of low school standing. Certain types of defects were found to prevail. They may be summarized as follows:

1. Inability to read typical material in an effective manner.
2. Study conceived largely as a process of memorizing the textbook.
3. Failure to organize and summarize the material studied.
4. Failure to review at appropriate intervals.
5. Lack of regular time and place for study.
6. Failure to concentrate upon the assignment.

7. The practice of quitting a lesson before completing the task to the best ability of the child.

Following the analysis of each student's difficulties, remedial instruction was given for a period. The results of this remedial training, while not successful in every detail, encouraged the authors to express the belief that if such study procedure could be carried out systematically during the entire school year, improvement would be evident in the majority of cases. Should this prediction prove true, it would mean the salvaging of many of our high school failures, resulting in a social and economic gain.

A brief examination of how bright children study by means of interviews seem to show conclusively that the success of such children was due, in part, to the efficient methods of studying their lessons. While the study is not an exhaustive one and the findings not entirely satisfactory, the study opens a new responsibility for the class-room teacher and supervisor.

The monograph closes with a bibliography of more than forty titles carefully annotated, thus making an important contribution for any one interested in a similar study.

J. CROSBY CHAPMAN AND GEORGE S. COUNTS. *Principles of Education.* Boston: Houghton Mifflin Company, 1924. Pp. 645.

The justification for this volume is found in a statement made in the introduction by the Editor: "There is danger that those in training today may grow up and pass out of our training institutions without gaining that sound grounding in the philosophy of the educative process which has been the great strength of the older generation of professional educators." The presence of such danger is certainly apparent to any one at all familiar with current educational literature and with the present tendency to emphasize "the newer engineering-type courses of instruction in education."

In organizing the subject-matter of the book, the authors use twenty-five important educational questions. Each of the larger questions is divided into a number of minor questions. This arrangement of the material lends itself admirably to classroom needs. The value of the book is further enlarged by including at the end of each chapter a serviceable collection of problems and exercises. Experience with the book warrants these statements.

The point of view of the authors cannot be expressed by stating the cardinal principles of any system of philosophy. It may be best presented in terms of what they conceive to be the large objective of education. *"This objective is that of inducting the child into the life of society and of training him in the use of the instrumentalities of civilization."* They believe that a sound philosophy of education must be founded upon human values and needs if it is to guide successfully a democratic society. Only when scientific procedure in education flows from such wholesome philosophy can the school advance toward its true goal. Such in brief is the point of view of the authors. While they have not given in detail the results of scientific investigations in the fields of psychology, sociology, and education, they have endeavored to incorporate in their book the spirit and fundamental truths of these researches. They have endeavored to express in a brief but a clear form the best that has been said regarding education from Plato to Wells. They have drawn from a wide group of authorities extending over the entire period of human history and representing the varied interpretations of life and its education.

Willis L. Gard,
Ohio University.

Fritz Wittels. *Sigmund Freud: His Personality, His Teaching, and His School.* Translated from the German by Eden and Cedar Paul. Dodd, Mead & Company. 287 pages. $3.50.

Herr Wittels, although he has been estranged from Sigmund Freud for years, is nevertheless an enthusiastic supporter of psychoanalysis and a thoroughgoing admirer of Freud. He speaks of Freud sometimes almost with awe. During the years of his estrangement he seems to have enshrouded the founder of psychoanalysis in a halo of romance. Freud to him is the Master, the overshadowing individuality, the Old Man of the primitive horde, a successor of Plato, one of those figures which "loom athwart the centuries." It is little wonder that Freud can write him the generally appreciative letter which forms the preface to the book. It is doubtful whether Herr Wittels, with such a profound admiration for Freud, can be the critical witness which he aspires to be.

His general attitude toward most of Freud's discoveries is illustrated by his enthusiasm for the Oedipus complex. The announcement of this complex, he says, "was acclaimed by those who were struck by the lightning of truth—and not by physicians alone, but by educationists, by serious-minded clerics, by sociologists, and, at length, by all who have a modicum of self-knowledge." The signs of the complex are so plain, in his opinion, "that it seems difficult to understand why the world had to wait until 1000 before a bold psychologist could discover this light for souls that have strayed from the path." These enthusiastic statements, however, bespeak only the general attitude of the author. He does not necessarily accept each and every pronouncement of Freud. For example, he does not answer Yes or No as to Freud's theory of the castration complex. The reader who is not an ardent Freudian may register a somewhat vigorous No—at least if he must depend for his evidence upon the information presented by Herr Wittels in his chapter on that subject.

The book, however, is valuable both for Freudians and non-Freudians. There is a vivid and entertaining account of the conflict of several strong personalities, among whom Freud is evidently the peer. The chapters on Adler, Jung, and Stekel show just where these well-known psychoanalysts broke with Freud, and just what each has contributed to psychoanalysis. There is also a clear account showing how Freud, step by step, built up his theories of the unconscious. The book, in other words, explains the inner workings of a movement which has become very powerful in our times.

B. L. JEFFERSON,
Ohio University.

MARION REX TRABUE. *Measuring the Results in Education.* American Book Company. Pp. 492.

Teachers have always tried to measure the results of their teaching, and the more progressive endeavor to use modern techniques. But the methods developed in the last few years by scientists have not been presented in an understandable way to the teacher. Either the scientist does not have the time, or he lacks the ability to interpret the results of his work so that the unskilled can use them.

Dr. Trabue happily combines the theory of scientific measurements with practical every-day problems which are familiar to the teacher in a way that can be understood by the average elementary school teacher. In the earlier chapters of the book he discusses the failure of school marks to measure effectively the results of education because of the lack of agreement of teachers in purpose, methods, and ability; the development of standard tests through the necessity of investigating and defending the conservatism of teachers; and the necessity of accuracy in securing the ages of pupils.

Statistical processes are begun in the fifth chapter, showing seriation, tabulations in graphic form in frequency tables and bar diagrams, and the calculation of the percentile points. Illustrations of types of measurement are used to emphasize the principles involved in each case. Tests for measuring spelling ability, the quality of handwriting, the speed and power in arithmetic, reading ability, and the quality in English composition, are discussed with clear directions for their use. In a clear graphic style the author shows the treatment of the data secured from test results, the calculation of measures of central tendency: the mean, the median, the mode; and measures of variability and dispersion: the mean, the median, the quartile, and the standard deviations; and the measurement of relationships: the coefficient of correlation.

It has been discovered that children differ greatly in their innate intellectual endowment, and this capacity, as well as the educational standing of the pupils, can now be measured by standardized tests constructed especially for the purpose. By the use of these tests we may discover many grades of intelligence, get helpful information in the classification of children, be able to give them better educational and vocational guidance, and more nearly adapt instruction to their capacities and needs.

The author advises the use of tests that have been scientifically constructed and standardized if the teacher wishes to have a basis of comparison with what other schools are doing, and also be able to compare quality and difficulty.

Sixty-three tables and fifty illustrations are valuable features of the book in making the data and treatment of test results clear and understandable.

C. J. PATTEN. *The Passing of the Phantoms.* A Study of Evolutionary Psychology and Morals. E. P. Dutton & Company, 1925. Pp. 103.

The author discusses rather sketchily the doctrine of recapitulation, and also the development of the moral sense through evolution.

The first chapter suggests some evidences of the reality of evolution as found in vestigial sex-organs, atavistic myology, the evolution of the nervous system from the skin-layer of the embryo in various forms of life along with manifestations in the early life of the babe of inherited race experiences; and that the explanation of the differentiation of the human embryo in its later stages from its cousins of all living beings depends upon "long associated habits of the cellular elements of the embryo itself."

In the second chapter the author reports from his observation of pets evidences of the evolution of mental powers. A Kestrel listened with rapt attention when sung to; and she had the imaginative power to "conjure up in her mind vague mental pictures of something animate . . . . which she had probably once upon a time seen and which frightened her," when the author approached her coop with a "hard black felt hat on . . . ." A dog was easily beguiled into the belief that harmless inanimate objects may possess "evil spirits." A horse stopped for a moment to gaze at a large sun-flower hanging over the hedge in the twilight, then bolted because there was in his mind a dim idea "of the presence of something uncanny." A tiger, disturbed over the sight of a mouse is reported to have entertained a mental state related to a fairy quadruped very similar to our superstition concerning fairies. And the tendency to imagine Spiritual Essence in natural objects began in creatures lower than man, Darwin's dog showed that the ethical sense was developing when he asked himself whether it was right to permit a strange "living" agent to move an open parasol on the lawn.

The author continues this line of reasoning in the third chapter to show the evolution of the moral sense from lower forms of life. Ants, which perhaps should rank next to man in the scale of intelligence, assist the feeble and are affectionate toward their fellows; gulls and terns exhibit sympathy for a wounded companion; and individual birds act as sentinels for the feeding flocks.

And there is no breach in continuity in evolution of morality from lower animals to man.

In the fourth chapter the author takes a fling at various theologies and dogmas, favors Agnosticism as an attitude (but does not adopt it), and finds that sympathy is the foundation stone for the moral sense. The evolution of the God-Devil ideas of man have come through the stages of animism, totemism, and astronomical myths. Many important myths and dogmas, such as the Resurrection of the Body, the Virgin Birth, of an Incarnate Deity in human form, had their origin chiefly in Solar Mythology. Those who believe in Super-Natural Beings are of the Superstitious Order, whose moral sense is permeated with "falsities and absurdities." But the non-superstitious man (few in number) is the salt of the earth because he accepts or rejects "statements according as his own reasoning faculty alone guides him."

There is a possibility of a case yet against evolution in the reader's mind, but not in the mind of the author.

J. R. GENTRY,
Ohio University.

TRUMAN LEE KELLEY. *Statistical Method.* New York: The Macmillan Company, 1923. Pp. xi + 390.

Yule's *Introduction to the Theory of Statistics* was published a decade and a half ago. During this period has taken place the greatest advance in the scientific aspects of education in its entire history. Hardly anyone taking an important part in this development has not acquired his knowledge of statistics to a considerable extent from a study of Yule's book. Indeed for more than a decade this work enjoyed a practically undisputed preëminence as a standard text in the subject. In 1923, however, appeared a work which seems destined to play for some years to come the rôle formerly played by Yule's *Introduction.* This new work is Truman Lee Kelley's *Statistical Method.*

In his introduction Kelley states that he designed the work chiefly for biologists, economists, educators, and psychologists. Probably the readers of this review will be interested in the work especially for its bearing on educational and psychological problems, particularly those concerned with tests. When Yule wrote his book this last great field of activity could hardly have been anticipated in its present state. Naturally many of the important statistical problems had not emerged. We find Yule's illustrations taken largely from economic and vital statistics. In gratifying contrast to this (from a psychological point of view) we find Kelley making frequent and pointed reference to statistical principles and devices of value both to the builders and the users of tests. As an example of this, on page 174 appears a table showing the estimating potency corresponding to various correlation coefficients. From the point of view of theoretical statistics there is nothing particularly novel about this table. From the practical point of view, however, it is extremely important. For example, if the estimating potencies of the various correlation coefficients as shown in this table had been thoroughly realized by the makers and users of tests during the last fifteen years, things might now be quite different. Many of the extravagant claims for tests and the equally extravagant expectations of them would have been impossible. On the other hand the equally indiscriminate condemnation of all tests, which has been partly the result of the over-optimism just mentioned, might also have been avoided.

One of the best things in the book is the masterful manner in which the author has treated the various implications flowing from a fundamental expression originally derived by Spearman. This largely makes up Chapter IX which is entitled "Functions Involving Correlated Measures." It is not too much to say that every person concerned with tests in any important way should be familiar with this chapter. Indeed the same might be said of the greater part of the book. It is high time that it be recognized that experimental work on tests is not a job for an amateur. On the contrary, it is distinctly a task for an expert, and one of the indispensable qualifications of such an expert must be a thorough knowledge of statistics. Kelley's *Statistical Method* will undoubtedly play a dominant part in the statistical training of the test expert for some time to come.

CLARK L. HULL,
University of Wisconsin.

A. J. SNOW. *Psychology in Business Relations.* Chicago: A. W. Shaw Company, 1925.

In this book the psychological problems of business are divided into five groups: Psychology of the Consumer, Psychology in Marketing, Psychology in Advertising, Psychology in Selling, Psychology in Employment.

The discussion of the consumer comprises a presentation of the traditional classifications of human conduct and a section written for the benefit of the consumer, in which are given a number of 'prophylactics' which may be used in withstanding the wiles of modern marketers. For example: "The consumer has opportunity to familiarize himself with goods by consulting other courses than the dealer alone. It is his privilege to know quality and price, just as much as the merchant's. His fellow-consumers have learned by experience, perhaps, from which he can profit, if he wishes. The man who has never owned an automobile, for instance, may profit by consulting his neighbor who does own a car before he makes his purchase." "The consumer's greatest defense, which cannot be overemphasized, is his knowledge of his own reactions in buying. By introspection, and by observing the reactions of fellow-consumers where the salesman uses certain methods, one may soon come to recognize the kinds of persuasion to which he ought to control his reactions. He must then learn to inhibit these reactions. Thus he guards not only against the tricks of the salesman, but also against his own impulsive actions."

The next three sections consist chiefly of general discussions of marketing situations, in great part condensed from the standard books on advertising and selling. A good deal of the material is of the 'commonsense' variety, such as: "It is important that the stock be as complete as possible." "The stock should be well located." "Good window and show-case display have important rôles. Feelings of pleasantness are produced through attractive displays of merchandise. Often desire for an article is induced by its setting in a show-case or window."

The fifth section deals with the difficulties involved in adjusting people to vocations; sketchy but clear treatment is accorded to the methods thus far developed for facilitating this adjustment: vocational analysis, tests, job specifications, and the like.

Publishers have long been vociferous in their appeals for a book on "Business Psychology." In view of the large number of courses of this title scheduled in the catalogs of Schools of Commerce and Colleges of Liberal Arts, one is tempted to regard these appeals as justified. One who tries to write such a book, however, encounters a number of perplexing questions: How much "general" and "theoretical" psychology to include; how much knowledge of an elementary sort can be expected from the reader; how can one write satisfactorily about the more subtle psychological relations in business without laying a considerable foundation of psychological theory? Then there is the circumstance that business relations are quite diverse; and while certain readers may be deeply interested in one phase like advertising, they may be uninterested in problems of employment; they can find the enlightenment they desire in the special books on advertising, so why should they buy a book on Business Psychology in general? The present volume is a valiant attempt to overcome these obstacles, and its reception will be watched with sympathetic interest.

The book calls forth several suggestions apropos of the writing of books of this character. Granting the real need of such summarized treatments of special phases of business, it would seem advantageous for an author to include some new material of his own. This will give value to the book even in the hands of those who do not care to use the more general sections. Dr. Snow has furnished some such new material, reporting some of his investigations on rating and on the measurement of taxi drivers.

In the second place there is much to be said in favor of reproducing tables and graphs from the work of other investigators rather than merely making abstract generalizations. Business psychology needs content. A respectable body of facts are accruing, and these should be furnished in their scientific dress to the more or less avid public, if for no other reason, in order to drive home the necessity of using scientific methods, and in order to illustrate these methods. In this respect, the book before us might be improved, mention having already been made of the generalized and summarized nature of most of the discussions.

The book is commendably reticent in the claims it makes for the potency of psychology in business and, providing it gives a clear enough picture of the psychological mechanisms involved, it will surely not mislead. Readers who desire, and are able, to consult technical reports of investigations in business psychology are given a well-selected list of references. For the instructor who wishes to assign class exercises, a series of problems are set at the end of each chapter, though, unfortunately (from the standpoint of the impecunious instructor) they require the purchase of a separate book of problems compiled by the author.

Mr. Paterson, unless he is by now thoroughly hardened, will wince at the customary misspelling of his name; as will also Mr. Rugg. The excellent piece of advertising copy of Miss White, which serves as frontispiece, should have been censored as to the spelling of "cereal." On page 416 "In the ideas . . . ." should probably read "If the ideas . . . ." Other minor corrections that would be made by a fastidious reader will probably be made by the author before a reprinting.

HARRY D. KITSON,
Indiana University.

TWO PSYCHOLOGISTS

The latter part of the nineteenth century saw the rapid development in America of a new science, the science of physiological or experimental psychology. The stimulus in this as in so much of our higher scholarship of that period came from Germany. Most of the leaders have now passed away. The two—G. Stanley Hall and Hugo Münsterberg—here considered were both trained under Wundt, were both influential in applied psychology and were both extraordinary examples of the union in the same persons of insight with facility of expression. Both were to a large degree popularizers, it would not be wrong to call them journalists. In one important respect they differed. Münsterberg, like James and Ladd, never broke with the philosophical tradition.

Hall did. Münsterberg tried to keep his idealism and his science in separate compartments. To Hall as to Comte philosophy was one of the stages through which the race has passed or is passing. Like Comte he nevertheless had a philosophy of his own.

G. STANLEY HALL. *Life and Confessions of a Psychologist.* New York: D. Appleton and Company, 1923. Pp. 623.

To G. Stanely Hall's own mind the most significant feature of his life as far as one may judge from this *Life and Confessions* was his rather unusual success in moulting the impenetrable shell of prejudice and ignorance in which life begins, his steady disillusionment, his progressive "eclaircissement," to borrow one of his own borrowed words. His work was the exorcism of dead faiths rather than the discovery of acceptable truths. He was too voracious for that. In his reading and in his intellectual hunger he was insatiable.

He had come far, this Puritan farm-lad who was now a leader of international standing in psychology and education. And yet his Puritan reserve persisted to the end and refused to let him reveal quite freely or to accept quite comfortably and naturally his religious position, his personal attitudes or the results of his own self-analysis. There is so often present in his revelations an element of strain, of labored bravura if not bravado. In this respect he is very modern indeed. "I am far older than my years," he writes (p. 596), "for I have laid aside more of the illusions and transcended more of the limitations with which I started than most. In the views I have attained of man, his place in nature, his origin and destiny, I believe I have become a riper product of the present stage of civilization than most of my contemporaries, have outgrown more superstitions, attained clearer insights, and have a deeper sense of peace with myself." It would not be easy to put such statements to any objective test. Elsewhere he says that all through life he was diffident before each new audience, that he was misunderstood, isolated and unable to be as friendly as he meant to be or to appear as genial as he really was. The laying aside of illusions gave pleasure but it also left scars.

Quite in line with the nature of his personal developments was the character of his influence in psychology. He was a purge and a great beginner but other men entered into his labors and made the positive contributions. And it is one of the poignant ironies of his life that of the many leaders in his field whom he inspired the most successful were those who diverged far from the paths and the topics that he approved.

Any man is free to reveal as much or as little of himself as he chooses or as he can. But repression at one point is liable to lead to over-expression at another. Was it not so here? Leaving the Williams of Mark Hopkins he spent a year at Union, associating the while with Beecher and he narrowly missed becoming a minister. Indeed he did become a prophet and the lines of his prophecy were at least partly laid down by the age in which he lived.

Although he was but a boy when the *Origin of Species* came out, that book had become the scientific bible by the time he reached Germany. Earlier or later the result might have been different. Now in New York by reading and in Germany by personal contact he came under the spell of Darwin, Haeckel, Strauss, Ludwig, Flechsig and, in particular, Wundt. Thus he became an evolutionary and genetic psychologist. Thereafter his major interests were two, psychology and education, if these be not, on his own interpretation, only one.

His student days in Germany over, Dr. Hall was in 1881 "surprised and delighted to receive an invitation from the Johns Hopkins University, then the cynosure of all aspiring young professors throughout the country to deliver a course of twelve semi-public lectures on psychology. The University at that time was adding another story to our educational system. . . . . (He) understood that this opened the ultimate possibility of a chair and spent the entire summer in the work of preparation," and was rewarded in 1882 with the appointment to a professorship. Before this he had already held positions at Harvard and Antioch. Now his real career opened and he found himself in charge of the first American laboratory of psychology. It is not necessary to quote again the list of remarkable men who studied with him here or to name over the experimental studies which issued from this institution under his direction. Certainly the introduction of experimental methods and of the Wundtian standpoint was one of the greatest services that any man has performed for psychology in America. It was here also that he established *The American Journal of Psychology*, "the first in its field in English." The interesting and in part amusing story of its founding is told (pp. 227 ff.) in the *Life and Confessions*.

The close of the Johns Hopkins period marks also the close of his experimental work in psychology. At Clark, Dr. Sanford was placed in charge of the laboratory and after the administrative difficulties had been partially overcome and after Dr. Hall returned to teaching he devoted himself in the main to evolutionary and genetic aspects of psychology. Psychogenesis is the name he gave to his basic course at Clark and the outline of it (pp. 363 ff.) shows how all-inclusive, pansophic, it was. But he justifies its character by saying that "interest, like steam in an engine, must be developed over a large surface, although when put to work it has to be applied to a small one." Whether this is good educational psychology would seem to depend upon the emphasis and the wisdom that are applied to the latter part of the statement. Undoubtedly his method brought out a large number of splendid workers in psychology and education; it would be interesting to inquire into the number that it spoiled.

Almost three hundred titles are required to cover the sheaf of work garnered by Dr. Hall in fifty years of professional activity. Even this is not a complete bibliography, of course. The larger number of these productions are articles, frequently famous articles like *The Study of a Sand Pile*, *A Study of Children's Collections* or *The Contents of Children's Minds*. About a dozen sizable volumes issued from his desk, two of them—the *Adolescence* and the *Educational Problems*—each containing approximately fourteen hundred pages.

It is probably not unfair to say that he had no systematic view in psychology. He proposed no closed scheme and he even seemed to react against exact methods. His interest was in growth and youth; and Youth he tried to serve. "My heart lives in the future and in this sense I am younger than youth itself, the nature of which I would chiefly understand and appeal to."

Everyone to whom the reviewer has submitted the question has assented to the tonic quality of the *Life and Confessions*. Is not this, his vigor and enthusiasm and youth, the secret of the greatness of G. Stanley Hall? The spade work had to be done. It was done by disciples. He would inspire.

MARGARET MÜNSTERBERG. *Hugo Munsterberg: His Life and Work.* New York: D. Appleton and Company, 1922. Pp. 449.

When Agassiz at thirty-nine came to America he had already made his great scientific reputation. During the remaining twenty-seven years of his prime he became a marvelously popular teacher and he founded a great museum but his scientific discoveries were not numerous. His eager audiences and his boundless opportunities for mere collecting distracted him from those intensive studies by which, chiefly, science is extended. There is some evidence, not indeed demonstrative but still significant, that America spoiled Agassiz as a scientist.

A like comparison may be made in regard to Münsterberg. He came earlier in his life, at twenty-nine, and without having a comparable amount of completed work behind him. But like G. Stanley Hall he had worked with Wundt, had been thoroughly trained in physiological psychology and his published work was full of promise and had received the enthusiastic approval of William James. Once settled in Quincy Street, Cambridge, opposite the large, old residence of Mrs. Louis Agassiz, his teaching schedule, the popular audiences always waiting to be interested but not always capable of receiving instruction, his fatal facility with voice and pen and a certain very human love of applause had their way with him; but not before he had done a creditable amount of some of the best work in psychology, pure and applied.

It is easy to fling the charge of superficiality. It is often supposed that one can not be both widely popular and profound. And this tradition has some virtue but there are exceptions and Münsterberg was one of them. He was both popular and profound although not often in the same book perhaps. The best of him is in the *Beitrage*, the *Grundzüge*, the *Eternal Values* and the laboratory studies, all of them scholarly and scientific works. The *Eternal Values* has been called "an idealist's confession of faith." It and the *Grundzüge* show the dualism and parallelism which were basic postulates of his world-view.

The breadth of his interests was remarkable. He was concerned with music, poetry and art, with international relations as well as with philosophy and psychology. It may be that it was because of this breadth that he understood as have few others the conditions and the limitations of scientific method. In the end, however, it would seem that he undertook too much. The curse of popularizing is this that once a man does very much of it he will not be allowed to do anything else. His fate resembles that of the literary artist who writes a funny book and whose public ever after compels him "to write as funny as he can." The story of his literary and platform activities reads like a romance. His later volumes were mostly propagandist in character. Such are *On the Witness Stand*, *Psychotherapy*, *Psychology and Industrial Efficiency*, *Vocation and Learning* and a host of articles in magazines and in the Sunday newspapers. Here he was exploiting the new psychology. Even here he was instrumental in furthering modern vocational psychology and helped to pave the way for the achievements of the psychologists in the army.

Münsterberg was not entirely fortunate in his day and generation at Harvard. He came in complacent 1892 when William James could calmly, and perhaps fairly, inform him that Harvard was the best university in America. Even in those days it is said that there was a group of Brahmans who did not understand him, refused to overlook his foibles and his vanities and who even tried to embarrass him. And the wild animosities of 1916–1917 worked a virtual crucifixion. At the examination of a graduate student for a higher degree a colleague and former friend, a leading American thinker, sat during the whole period next to Münsterberg without a sign of recognition. Probably no one any longer believes him to have been a political agent but his career shows how hard it is to be friendly to both of two warring nations.

Perhaps this book also would be more convincing if it did not so clearly try to make out a favorable case for its subject. Biographies made in the family almost invariably suffer from this fault. But the volume is well-written, it presents the necessary facts for an understanding of Münsterberg's career and in the appendix there is an adequate summary of his psychological and philosophical writings.

H. G. Good,
Ohio University.

Robert H. Lowie. *Primitive Religion.* New York: Boni and Liveright, 1924. 346 pp. $3.50.

Primitive Religion is a comparative study of primitive religion, psychological in point of view, and undertaken for the purpose of finding the "least common denominator in all religious phenomena." Professor Lowie defines religion as a "sense of something transcending the expected or natural, a sense of the Extraordinary, Mysterious, or Supernatural," yet he would have it understood that this is "but a minimum definition of the psychological correlates of religion. . . . ." Following Drs. Marett and Goldenweiser, the author discovers the sense of the supernatural as the fundamental, common minimal element in religion, "not in the arbitrary division of the Sacred from the Profane but in the differential response to normal and abnormal stimuli, in the *spontaneous* distinction thus created between Natural and Supernatural, which does not require any pre-existing abstract formulation of 'nature.' The response is that of amazement and awe; and its source is the Supernatural, Extraordinary, Weird, Sacred, Holy, Divine." In a word, that is labelled religious which in the subject is an attitude of wonder, amazement, fear and awe, and which surrounds its objective correlates, the stimuli, with a strong affective element involving the claims of reverence, respect or worship.

This definition is based chiefly on the study of the psychological content of four contemporary primitive religions which are believed to be representative: the Crow Indians of eastern Montana, the Ekoi tribe on the west coast of Africa, the Bukaua with their neighbors, the Jabim and Tami of New Guinea, and the Polynesians. Having acquired an intimate knowledge of the beliefs and practices of these tribes, the author seeks to interpret these beliefs and practices as the votaries themselves view them. Part II, a *Critique of Theories*, is a critical comparison and evaluation of the author's conception of religion with the Animism of Tylor, the Magic of Frazer, and the Collectivism of Durkheim. Part III, headed *Historical and Psychological Aspects*, is a somewhat disconnected group of essays in which the author further seeks to justify his theory.

The reviewer is much in doubt whether Professor Lowie's source material for the development of even a minimal definition of religion is sufficiently diverse and representative. Why limit the source material to a few contemporary primitive religions? Why not include past religions as well as present? and why not include the history of a few great religions as well as religious phenomena as found in different stages of human civilization and culture? The language, too, in many cases is unnecessarily technical, certainly beyond the powers of a beginner in comparative religion. One is favorably impressed, however, with the author's scholarly treatment of the religions selected. This is his real contribution to comparative religion. Extremely pertinent is Professor Lowie's last word which is an emphatic caution (1) to those who would devise their definition of religion by *a priori* methods without factual data gained through painstaking observation and research; (2) to those who in their zeal for simplification of the diverse religious phenomena tend to overlook or ignore important individual differences; and (3) to those who criticise any definition of religion before inquiring into the point of view or purpose which has served as an organizing principle. Perhaps there is no single adequate definition of religion.

WALTER S. GAMERTSFELDER,
Ohio University.

NEW BOOKS AND PAMPHLETS RECEIVED[1]

Books and pamphlets for review should be sent to James P. Porter, Department of Psychology, Ohio University, Athens, Ohio.

Clinical Psychology. Louis E. Bisch. The Williams and Wilkins Company, Baltimore, Md. Price $3.00. 340 pp.

Collected Papers of Sigm. Freud. Vol. III. The Hogarth Press, 52 Tavistock Square, London, W. C. 1. Pp. 605.

Constitution at a Glance, The. Henry B. Hazard and Margaret D. Moore. Henry B. Hazard, Lock Box 1919, Washington, D. C. Price $.75.

Crisis in Psychology. Hans Driesch. Princeton University Press, Princeton, N. J. Price $2.50. 275 pp.

Education as the Psychologist Sees It. W. B. Pillsbury. The Macmillan Company, New York, 1925. Pp. 342.

Improving Rural School Instruction and Supervision in Colorado. Joseph H. Shrider and L. Thomas Hopkins. University of Colorado Bulletin, Vol. XXV, No. 1. Boulder, Colorado. Pp. 99.

Linking Science and Industry. H. S. Person, E. K. Hall and others; edited by Henry C. Metcalf. The Williams and Wilkins Company, Baltimore, Md. Price $3.50. 206 pp.

Mental Growth of Children. Buford J. Johnson. E. P. Dutton and Company, New York. Price $3.00. 160 pp.

l'Odorat: Encyclopedie Scientifique. H. Zwaardemaker. Gaston Doin, Editeur, 8, Place de l'Odeon, Paris. 305 pp.

Primer of Graphics and Statistics for Teachers. Harold Rugg. Houghton Mifflin Company, New York. Price $1.60. 142 pp.

Professional Secondary Education in Teachers Colleges. Alfred Lawrence Hall-Quest. Columbia University Contributions to Education, No. 169. Columbia University, New York.

Program for Higher Education in the Church of the Brethren. John Samuel Noffsinger. Columbia University Contributions to Education, No. 172. Columbia University, New York City.

Psychology in Advertising. Albert T. Poffenberger. A. W. Shaw Company, New York. 632 pp.

---

[1] Mention here does not preclude further comment.

Report of Statistics of Manufacturers for 1921: Massachusetts. The Secretary of the Commonwealth, State House, Boston 9, Mass. 100 pp.

Social Psychology. KNIGHT DUNLAP. The Williams and Wilkins Company, Baltimore, Md. Price $4.00. 261 pp.

A Study in the Psychology of Learning in Geometry. WINONA M. PERRY. Columbia University Contributions to Education, No. 179. Columbia University, New York.

A Study of Learning and Retention in Young Children. LOIS HAYDEN MEEK. Columbia University Contributions to Education, No. 164. Columbia University, New York.

Trancendental Values. R. P. BYERS. Richard G. Badger, Publisher, Boston, Mass. 60 pp.

PUBLICATIONS ISSUED BY THE BUREAU OF EDUCATION, DEPARTMENT OF THE INTERIOR, WASHINGTON, D. C.

Courses in Rural Education Offered in Universities, Colleges, and Normal Schools. KATHERINE M. COOK. Rural School Leaflet No. 37, March, 1925. Price 5 cents. 18 pp.

High School Education of the Farm Population in Selected States. E. E. WINDES. Bulletin No. 6, 1925. 24 pp.

Important State Laws Relating to Education, 1922–1923. WILLIAM R. HOOD. Bulletin No. 2, 1925.

Kindergarten Legislation. NINA C. VANDEWALKER. Bulletin No. 7, 1925. 32 pp.

Per Capita Costs in City Schools, 1923–1924. FRANK M. PHILLIPS. Statistical Circular No. 4, March, 1925. 7 pp.

Record of Current Educational Publications. JOHN D. WOLCOTT. Bulletin No. 14, 1925. 50 pp.

Statistics of City School Systems, 1921–1922. FRANK M. PHILLIPS. Bulletin No. 34, 1924. 222 pp.

Visual Education and the St. Louis School Museum. CARL G. RATHMANN. Bulletin No. 39, 1924. 36 pp.

EMPIRICAL DATA ON THE SCORING OF TRUE-FALSE TESTS

DONALD G. PATERSON AND T. A. LANGLIE

University of Minnesota

In a recently compiled bibliography of fifty-six references to new type examinations, twenty related to or discussed the true-false type of examination.[1] Eight of these references contained experimental evidence concerning the reliability and validity of true-false tests utilizing the traditional right minus wrong scoring formula; eleven advocated or severely criticized, on theoretical grounds, the use of true-false tests and the right minus wrong scoring method in connection therewith; only one reference presented experimental evidence regarding the comparative reliability of true-false tests when scored right minus wrong or when scored number right.[2]

In all of this work, psychologists have proceeded to utilize the true-false test on the basis of the following assumptions, explicit or implied:

a. Guessing enters into true-false tests to a marked extent.

b. The effect of guessing should be and is counteracted by a right minus wrong scoring formula.

c. The right minus wrong scoring method is more reliable than merely scoring by the number right or by the number wrong.

PREVIOUS WORK ON BEST SCORING METHODS FOR TRUE-FALSE TESTS

Ruch's discussion of these assumptions is penetrating, leading him to secure experimental data which throw serious doubt on the truth of these assumptions.[3] His data were derived from duplicate forms of two-alternative response tests in social science and history and also from duplicate forms of the three two-alternative response tests in the Terman Group Test for Mental Ability. A summary of his results follow:

---

[1] Donald G. Paterson, The Preparation and Use of New Type Examinations, Appendix B, Annotated Bibliography, Published by the World Book Company, New York, 1925.

[2] G. M. Ruch, The Improvement of Written Examinations, ix, pp. 1–193, Published by Scott, Foresman Company, New York, 1924.

TEST	RELIABILITY FOR DUPLICATE FORMS OF 50 ITEMS EACH	
	Scored number right	Scored right minus wrong
Ruch, 2-response (135 cases)	+0.74±0.027	+0.68±0.031
Ruch, true-false (133 cases)	+0.56±0.040	+0.41±0.049
Terman, Test 3, Word Meaning, 2-response (43 cases)	+0.67±0.056	+0.56±0.071
Terman, Test 6, Sentence Meaning, Yes-No. (43 cases)	+0.53±0.073	+0.47±0.080
Terman, Test 8, Mixed Sentences, True-False (43 cases)	+0.68±0.056	+0.56±0.071

In each and every instance the number right method of scoring results in higher reliability coefficients than the right minus wrong method.

The writers are aware of only one other reference making experimental comparisons of these two methods of scoring, i.e., analyses made in determining the best methods of scoring the army intelligence tests.[4] In discussing "examination a" the army report states:

> The tests were scored in various ways in order to determine the best method empirically. The principle adopted was that of finding by trial the method of scoring which would give the highest correlation with Binet or with total scores. All the tests were scored both for "number right" and "right minus wrong." In addition, Test 4 was also scored for "right − 2 wrong"; and Tests 5 and 7 for "3 right − wrong." It has been supposed that it would be found necessary to penalize, more or less extensively, for errors. However, it was found that, in general, penalizing was of little or no advantage except in those tests where it was

---

[3] G. M. Ruch, op. cit., pp. 114–120.

[4] Memoirs of the National Academy of Sciences, Vol. XV, 1921.

> necessary to counterbalance the factor of chance, as in Tests 3 and 6. Test 4, for example, gave significantly higher correlations with Binet and with total score when scored for number right than when scored for "right − wrong" or "right − 2 wrong." Only in Tests 3, 6, 7, and 10 did the data justify subtraction from the score in case of error. In Tests 3 and 6 the subject has an even chance of making a correct response by guessing, and this is offset by scoring the test "right − wrong." In Tests 7 and 10 the advantage gained by penalizing for errors was small.[5]

One is unable to verify the soundness of the above conclusions because no data, in the form of number of cases, means, sigmas, reliability coefficients, or probable errors are given.

In a section devoted to methods of scoring "examination a," the same problem is presented with more details.[6] A study made at Camp Lee for a group of random cases revealed the following for the 2 response tests:

	CORRELATIONS WITH TOTAL EXAMINATION SCORES	
	Number right	Right minus wrong
Test 3, True-false	+0.66±0.028	+0.72±0.024
Test 6, Same-opposites	+0.76±0.034	+0.82±0.026

Here, we find a slight advantage for the right minus wrong method. However, the difference in favor of the right minus wrong method is not statistically significant, i.e., the difference in neither case is equal to 4 or even 3 times the probable error of the difference. These results, it is to be noted, are not in terms of reliability but in terms of validity according to current usage. The importance of this asserted advantage is lessened by the next sentence, "Similar data collected at Camp Dix indicated a slight advantage for the right minus wrong method for test 3 and no advantage for this method in the case of test 6." The above constitutes the sole evidence for the adoption of right minus wrong methods in scoring these two tests which became tests 5 and 4 in the Alpha examination.

---

[5] Memoirs , op. cit., p. 305.

[6] Memoirs, op. cit., p. 339.

RELIABILITY COEFFICIENTS FOR DIFFERENT SCORING METHODS

The above presentation of available evidence indicates the need for additional data in order that this question of the best method of scoring two-alternative response tests may be more definitely established. The writers, therefore, present the following data to be considered in connection with the preceding results and to serve as a stimulus for additional contributions of a like nature in various school subjects. The results here reported were obtained from a 100 item true-false test (unlimited time allowed) covering twelve weeks' work in the Psychology of Advertising, involving 61 juniors and seniors in 1924 and 50 juniors and seniors in 1925. The students were cautioned against guessing and informed that a right minus wrong scoring formula would be used in grading

TABLE 1

Reliability coefficients for 100 item true-false test in psychology of advertising according to the method of 50 odd items vs. 50 even items

SCORING METHOD	CLASS IN 1924 (61 CASES)	CLASS IN 1925 (50 CASES)	BOTH CLASSES COMBINED (111 CASES)
Number right................	+0.57±0.060	+0.73±0.045	+0.63±0.037
Right minus wrong..........	+0.49±0.065	+0.60±0.061	+0.54±0.045
Number wrong..........	+0.46±0.069	+0.57±0.063	+0.51±0.048

TABLE 2

Statistical constants for groups in table 1

	CLASS IN 1924		CLASS IN 1925		BOTH CLASSES COMBINED	
	M.	S.D.	M.	S.D.	M.	S.D.
Number right, odds...........	41.0	3.6	40.8	3.7	41.4	3.7
Number right, evens..........	43.1	3.5	42.1	3.1	42.6	3.4
Right-wrong, odds.............	35.2	5.9	33.9	6.0	34.6	6.0
Right-wrong, evens...........	36.9	5.9	36.1	5.5	36.6	5.8
Number wrong, odds..........	7.3	2.7	7.4	3.0	7.3	2.9
Number wrong, evens.........	6.6	2.6	6.2	2.8	6.4	2.7

their papers. Reliability coefficients were derived by correlating 50 odd items vs. 50 even items.[7] This method, while not the best method of measuring reliability or consistency, is not open to the charge of spuriously magnifying consistency which frequently results when odd items are correlated with even items in *time limit tests*. The reliability coefficients are summarized in table 1.

The means and standard deviations of the distributions for the above groups are presented in table 2. Inspection of table 1 reveals the fact that for both classes combined and for each class separately the number right scoring method yields the highest reliability coefficients, the right minus wrong method yields the next highest and the number wrong method yields the lowest reliability coefficients. The differences, even in the case of the combined classes, are not statistically significant yet they duplicate Ruch's results. Hence, the assumption that the right minus wrong method is more reliable than the number right method of scoring true-false tests is seriously questioned by the facts. Being unable to explain the differences in the reliability coefficients for the 1924 class and the 1925 class, the writers have included the means and sigmas in table 2. In no case are the differences statistically significant, the inference being that 50 item true-false tests with groups of fifty or sixty students will yield fluctuating reliability coefficients from group to group. A difference of as much as 16 points may occur and still be accounted for on the basis of chance variations in the repeated experimental conditions.

VALIDITY COEFFICIENTS FOR DIFFERENT SCORING METHODS

In determining the validity of the different scoring methods, average grades in all examinations in the course constituted the criterion, the results of the true-false tests themselves being omitted. In 1924 the criterion was a combination of essay and objective type examinations. In 1925 objective type examinations only were used, consisting of completion and recognition items. In combining average scholarship for the two classes the mean of each was made equal to 50, tenths of sigma being counted as 1 each. A scholarship score of +1 sigma thus becomes 60 and a scholarship score of −1 sigma thus becomes 40. The assumptions involved in this procedure are: (1) the two classes are of equal ability on the average, (2) the variability in each class is the same, and (3) the distribution of scholarship in both classes conforms to the normal curve. There is reason to believe that these assumptions are sufficiently correct to warrant the combination of scores in the manner indicated.

---

[7] All correlations reported in this paper were computed by the Pearson Product Moment formula, each and every step being double checked. The reliability coefficients in each instance were computed from class intervals of 1 each, thus avoiding any possible errors due to grouping scores in class intervals.

TABLE 3

Correlation between each scoring method for 100 item true-false test and average scholarship

SCORING METHOD	CLASS OF 1924 (61 CASES)	CLASS OF 1925 (50 CASES)	BOTH CLASSES COMBINED (111 CASES)
Right minus wrong.........	+0.39±0.073	+0.52±0.070	+0.44±0.050
Number right...............	+0.38±0.074	+0.49±0.072	+0.39±0.054
Number wrong.............	−0.35±0.076	−0.52±0.070	−0.44±0.051

TABLE 4

Statistical constants for variables in table 3

	CLASS OF 1924 (61 CASES)		CLASS OF 1925 (50 CASES)		BOTH CLASSES COMBINED (111 CASES)	
	M.	S.D.	M.	S.D.	M.	S.D.
Average scholarship (criterion)	50.4	9.7	53.1	10.1	51.6	9.9
Right minus wrong............	71.4	10.1	69.8	10.1	70.7	10.1
Number right.................	84.7	6.0	82.4	6.4	83.6	6.3
Number wrong................	13.4	4.7	13.0	5.2	13.3	4.9

Table 3 presents the validity coefficients and table 4 presents the means and standard deviations of the variables involved.

Reference to table 3 shows slight differences between the validity coefficients for the various scoring methods. Such differences as do exist are not statistically significant. The data, although indicating the possibility of a slightly higher validity for the right minus wrong method, fail to substantiate the oft repeated claims for the necessity and desirability of the right minus wrong scoring formula. Indeed, here we have evidence that merely scoring the number wrong is as valid as making the correction for chance although no one has ever seriously proposed such a method of scoring true-false tests.

Inspection of table 4 shows that the correction for chance (right minus wrong method) lowers the mean scores in the 100 item test to a marked extent, at the same time greatly increasing the standard deviations. This greater variability would be a decided gain were it accompanied by increased reliability, because the corrected scores would then magnify the individual differences within the group. The magnification of individual differences is not genuine, however, because they are not necessarily true differences, the reliability of corrected scores actually being less than the reliability of uncorrected scores. It is possible that this increased variability is responsible for the more favorable showing of the right minus wrong scoring method with reference to validity.

ADDITIONAL RELIABILITY COEFFICIENTS

The data presented by Ruch and that presented in this paper are faulty in that the number of cases involved is rather small. For this reason, the writers obtained test blanks involving two duplicate forms of a same-opposites test given to 519 University of Minnesota sophomores by R. C. McKee in 1922–1923.[8] These two forms of the same-opposites tests were obtained by selecting two lists of 50 stimulus words from a list of 250 words which had been standardized to determine the difficulty of each. These two forms, A and B, were scaled according to P.E. difficulty and given to 519 students who were instructed to guess. A short time limit was also imposed. McKee scored them according to the right minus wrong formula saying, "This procedure of scoring was used because it was possible to obtain a score of 50 per cent correct by chance. (See W. A. McCall, How to Measure in Education, pp. 127–128, for a full discussion of this scoring device.)" To utilize McKee's data for our present purposes necessitated rescoring all of the blanks, thus obtaining three scores for each student on Form A and on Form B. The three scores were the number right, the number wrong and the number right minus the number wrong. Table 5 presents the reliability coefficients computed for each of these methods of scoring, together with the means and standard deviations.

---

[8] R. C. McKee, A Study of Controlled Association Technique. Unpublished M.A. Thesis deposited in the University of Minnesota Library, 1923.

TABLE 5

Reliability coefficients for a same-opposites test according to the method of correlating scores in Form A with scores in equivalent Form B.

Number of cases equals 519

SCORING METHOD	r_{ab}	FORM A		FORM B	
		Mean	S.D.	Mean	S.D.
Number right	+0.69±0.016	38.0	5.5	38.0	5.5
Right minus wrong	+0.63±0.018	34.6	6.0	34.6	6.1
Number wrong	+0.43±0.024	3.9	2.5	4.0	2.7

Again, we note that the number right method of scoring yields a higher reliability coefficient than either the right minus wrong method or the number wrong method. The difference, as before, is not statistically significant even though the large number of cases makes the probable errors of the *r's* very small. The difference is actually less than three times the P.E. of the difference.

SUMMARY AND CONCLUSIONS

1. The reliability of a 100-item true-false test in two Psychology of Advertising classes was actually lessened when

the right minus wrong scoring formula was used to counteract the effect of guessing. The number right method was most reliable, the right minus wrong method was next, and the number wrong method was least reliable, although the differences are not statistically significant. These results hold for each class separately and for both classes combined, and are in direct agreement with Ruch's experimental findings.

2. Regardless of method of scoring, the reliability coefficients are low, throwing doubt on the true-false test as a universally reliable examining device.

3. Correction for chance by the right minus wrong method as compared with simply scoring the number right lowers the mean scores and increases the variability. This larger variability is not accompanied by higher reliability, hence it does not constitute any real gain.

4. The validity coefficients show no statistically significant differences. Though such differences as happen to exist seem to favor the right minus wrong method, yet scoring by number wrong is equally valid.

5. The validity coefficients, regardless of method of scoring, are low leading to the tentative conclusion that the true-false test is not universally a highly valid method of examination.

6. Additional data, giving reliability coefficients based upon the testing of 519 students with two duplicate forms of a same-opposites test, show that the number right method of scoring yields a higher reliability coefficient than either scoring right minus wrong or number wrong, although the differences are not statistically significant.

7. The above results corroborate Ruch's evidence at every point. No statistically significant difference appears in favor of the right minus wrong method of scoring. Indeed, the number right method actually produces higher reliability coefficients in every instance. It is more reliable regardless of whether the two-alternative response tests are given with or without a time limit, and regardless of whether the students are told to guess or are warned not to guess. These results are contrary to the assumptions which psychologists have so

confidently made in urging a right minus wrong scoring method in conjunction with two-alternative response tests.[9]

8. Incidentally, these investigations reveal the unreliability of the true-false, yes-no, or same-opposites test when only 50 items are used. Brown's prophecy formula applied to the data in this paper would show that at least 100 items need be used to produce a passably reliable two-alternative response test and that 200 items need be used to produce a thoroughly reliable test. Single-word-answer questions, completions, analogies and matching or pairing questions can be made to produce reliabilities of +0.85 to +0.90 much more economically. The writers predict that the two-alternative response test, with the right minus wrong dogma destroyed and with its reliability demonstrated to be frequently low, will be forced to yield to these other forms of new type examinations.

[9] One of the writers, in publishing material in advocacy of new type examining procedure, has been equally guilty of theoretically defending corrections for guessing without any such empirical evidence as is here presented. Henceforth, he will not insist on a right minus wrong formula and indeed will not even advocate the use of true-false tests unless they are multiplied to uneconomical limits. See Donald G. Paterson, Improving the Examination Function in Teaching, Bulletin of the University of Minnesota, Vol. XXVI, No. 31, August 4, 1923, pp. 47–56.

THE STATISTICAL VIEWPOINT IN VOCATIONAL SELECTION

MAX FREYD

Personnel Research Federation, New York City

In a recent number of the JOURNAL OF APPLIED PSYCHOLOGY Viteles makes a plea for the clinical viewpoint in vocational selection.[1] He expresses the opinion that the statistical viewpoint is the representative one among research workers in selection and that this viewpoint, while appropriate to a technician, is not completely psychological. The points which he raises are of fundamental importance to experimenters in this field and deserve careful consideration by them.

The psychologist just as any scientist endeavors to collect a body of related facts established by scientific method. He relies essentially upon experiment to enlarge his body of knowledge. Because his science is comparatively young and has been subject to disputes on method the number of established facts is still not large.

Particularly in vocational selection where psychology made a rather late start have results been slow to accumulate. Nevertheless, by the use of the experimental method facts on vocational selection are gradually increasing. There is no other method which psychology can use to increase its body of knowledge in this field.

Let us review briefly the experimental approach to a problem in selection, in order to see the manner in which the statistical viewpoint operates.

The psychologist is usually called in when the employment department has failed to solve a definite problem such as the selection of the right men for a certain job. The management has heard that psychology is the science of human nature and what is more reasonable than for it to suppose that such a science can set it right. The system of hit-or-miss guesswork has failed and science is called upon to take its place.

---

[1] M. S. Viteles, The Clinical Viewpoint in Vocational Selection. Journal of Applied Psychology, 1925, Vol. IX, 131–138.

Even the best psychologist cannot sit down at once and pick the right men for the job. More likely than not he knows nothing about the work. He does not know what abilities are necessary to succeed in it nor even what constitutes success in it. In consultation with the management he decides carefully upon a criterion of success in the job and selects excellent, fair, and poor workers for study. He analyzes the job carefully from every angle, observing the employees, and perhaps doing the work himself. He gradually accumulates material on the psychological, social, and economic aspects of the occupation, all with reference to the requirements for success in it.

The completion of this step still does not supply him with sufficient means for the selection of applicants. He must have some way of measuring the applicants with regard to the requirements for the job. For this purpose he constructs tests, rating scales, and questionnaires.

At this point the employment psychologist has two uncertainties. He must have doubts about the accuracy of his analysis of the requirements for the job and about the validity of his tests as measures of these requirements. He resolves both doubts by giving the tests to workers on whom he has some criterion of vocational success. He compares test scores with vocational success by statistical methods, minimizing the effect of disturbing variables, providing sufficient cases, and computing the error of his results. If he finds that a test correlates highly with vocational success he assumes that the close causal connection between the two variables shows that they call out the same abilities. Whether or not his original analysis of abilities was correct is not important to him now. Nor is it important to him whether or not his tests measure those particular abilities for which they were constructed. The tests certainly measure some of the abilities necessary to success and his original analysis proved to be valuable in suggesting the kinds of tests to use.

The fact that careful statistical methods disclose a high relationship between test score and criterion warrants the use of the tests in selection for the job. Among applicants the known variable will be the test score and the predicted variable will be success at the job.

All assumed factors in success other than those represented in the tests are verified experimentally by rating scales and questionnaires and with the use of the same statistical methods as in the case of the tests.

Since these forms are standardized and quantitative they may be given by a trained examiner. He must be tactful and polite so that the worker is not led to believe that he is being put through a mill. The psychologist's further duties are to train the examiner, make certain from time to time that the examining methods are carried out as directed, and conduct further research with the purpose of bettering these methods and adapting them to changes in the conditions surrounding the job.

This in barest outline shows how the statistical viewpoint operates. It tries to put the selective process on a firm scientific basis. It tries to eliminate guesswork by providing objective proof to back every judgment. The examiner does not assume that he has any way of improving the selecting methods other than by experimental investigation.

The clinical point of view agrees with this up to a certain point. To quote from the article referred to:

> It is the opinion of the writer that in the cause of greater scientific accuracy in vocational selection in industry the statistical point of view must be supplemented by a clinical point of view. It must be recognized that the competency of the applicant for a great many jobs in industry, perhaps even for a majority of them, cannot be observed from an objective score any more than the ability of a child to profit from one or another kind of educational treatment can be observed from such a score. There is no reason for suspecting that the capacity of an individual motorman to avoid accidents, or of a printer's apprentice to profit

> from instruction in this trade can be expressed in an objective score, as easily interpreted by a minor clerk as by a trained psychologist, than for suspecting that the mental status of a child is revealed in the I.Q. which can be obtained by any teacher who owns a copy of Terman's Condensed Guide and a set of testing material. The one problem is as complicated as the other; the objective score in one case has in it as many elements of error as in the other, and an adequate diagnosis in both involves interpretation by a trained psychologist based on observation of performance and a consideration of related data. . . . .
>
> But industry must also be told with equal frankness that there are a great many jobs for which workers can not be adequately selected by tests administered and interpreted by employment clerks. It must be told that the selection of applicants for such jobs involves an examination by a trained psychologist who depends upon his scientific knowledge of human behavior as well as upon the test results. . . . .

This appears very much as if it were a case of changing horses in the middle of the stream. The psychologist starts out to measure the abilities required for the job. When he has evaluated statistically a series of tests he states that the tests are inadequate for the purpose and that the applicants can be selected properly only by a supplementary judgment of their performance which the psychologist alone by virtue of his special training is competent to give.

The field for the psychologist in industry is experimental research. If other fields seem greener they are a mirage. As soon as the psychologist forsakes his experimental results for his personal judgments he comes into competition with shrewd business men who have not had the benefit of laboratory training but have been forced into contact with all types and sorts of human beings. Is there any reason to believe that a psychologist can select men more accurately by adding his "scientific knowledge of human behavior" than a capable and experienced employment manager who is also in possession of the test results and is aware of their significance? The writer knows of no experimental data in support of this belief. There is plenty of evidence to show that psychologists do not agree among themselves in observing and judging human abilities even in employment procedure.

Let us see where the clinical viewpoint leads. If the psychologist uses his trained judgment to interpret and supplement the tests and fails to select men accurately the results will be unfortunate to say the least. He will demonstrate his own inability to judge men correctly and will discredit the science as well. If he succeeds correct psychological procedure demands that he be guided by the statistical viewpoint.

If he succeeds in selecting men efficiently he either can or cannot point to the evidence upon which he based his correct decisions. If he cannot point to the facts then he has judged not as a psychologist but as a shrewd layman and neither the industry nor the science have gained by it.

Let us assume that he does point to certain evidence, other than test scores upon which he based his correct judgments. He observed that certain applicants were extremely nervous and upset while taking the tests and were barely able to restrain themselves from flying off the handle. He states that the observation of this fact as well as the test score accounted for his success in selecting men. If another made such an assertion the psychologist would be the first to demand proof of the validity of the procedure. He has shown exactly this attitude toward Dr. Blackford. While admitting her ability to judge men he has questioned her statement of the points on which she judges them.

The psychologist cannot point to the factors other than test scores upon which he based his correct judgments unless he keeps a record of his objective judgments on these factors and compares these records with the vocational success of the men judged. Thus he is forced to adopt the statistical viewpoint.

The psychologist's methods of judging these factors objectively can be incorporated into a rating scale and taught to a bright and sensible employment manager or examiner. The examiner does not have to know the entire field of psychology. He does have to know the technique in giving and scoring the particular tests and forms and the quantitative or qualitative observation of the other variables which the psychologist has proved to be important in selection.

If this procedure is carried out psychology is the gainer because facts are established and other research workers share information which otherwise would exist uncrystallized in the brain of one psychologist.

There is another form of judgment in selection where the clinical viewpoint at first glance seems to have a place. If the standings of the applicant in all the prognostic agencies are plotted the curve is commonly called the mental profile of the applicant. Presumably it becomes the function of the psychologist to diagnose the applicant's fitness for the job by observing this profile and weighing one ability against another. Anyone who has tried this method knows how difficult it is. There are innumerable kinds of profiles and in many cases the judgment of fitness is the wildest kind of guess. When statistical methods such as total correlation, critical scores, and weighting are in existence there is no justification for diagnosing mental profiles.

One of the assumptions of the statistical point of view, according to Viteles, is that of the relative unimportance of the individual worker in industry.

> A selective process which is satisfied with the probable adequate selection of a group of workers, rather than with the proper placement of the individual worker (and such is the case in selection based upon objective scores alone) fails to give adequate consideration to the well-being and interest of the individual worker, an interest which in comparison with the point of view of European psychologists, has been shamefully neglected by the American psychologist.

It seems to the writer that the statistical viewpoint does not suffer by comparison here. The discussion concerns vocational selection and not vocational guidance. The employment office is aware of every vacancy in the company and can test the applicant's fitness for each of them. If he fails to measure up to the standards of any of the jobs the company cannot be interested in him any further. The clinical viewpoint can add nothing to this. In fact it may be ventured that the clinical viewpoint by allowing selection to be influenced by personal interpretations with their unavoidable prejudices instead of relying upon objective measures gives even less consideration to the well-being and interest of the individual worker.

The writer is not in disagreement with the clinical point of view if it means that social, personal history, and economic information should receive the same consideration as psychological abilities. These very important facts have been unaccountably neglected by research workers in spite of the fact that in many instances of their use, notably with life insurance salesmen, they have proved to be important. When they are studied they should be evaluated in the same statistical way as the tests. They can then be incorporated into questionnaires and rating scales and determined by the trained examiner.

There are some fields of psychological activity where the clinical method is justified. Vocational guidance as distinct from vocational selection is one of these. The difficulties in the way of checking the accuracy of vocational guidance are almost insurmountable. Progress must wait on the establishment of scientific facts in vocational selection in order that these may be applied in guidance. Meanwhile there are some proved measures as well as social and economic facts for the use of the psychologist and for the rest it may be granted that his advice will be as trustworthy as any other influence in motivating the youth to go into one or another vocation. But vocational selection offers a comparatively clear experimental problem with an opportunity for verification of judgments and when this opportunity exists there is no justification for adhering continuously to the clinical point of view.

The same writer upholds the clinical viewpoint in another recent paper.[2] Here he presents a number of principles "fundamental to the use of tests in the adequate guidance of the individual." We are interested in only four of these principles which are given in the author's words.

> *(4) The factors which determine vocational competency are the individual specific mental abilities. For positive guidance the specific abilities of importance in the various occupations must be subjected to measurement.*
>
> . . . .
>
> *(5) Tests, to be of service in guidance, must measure not only the general*

---

[2] M. S. Viteles, Psychological Tests in Guidance: Their Use and Abuse, School and Society, 1925, Vol. XXII, 350–356.

> *and specific mental abilities which are of importance in vocational success; tests must also be provided to measure the temperamental and character qualities which play so great a part in vocational adjustment.* . . . .
>
> (8) *In the examination of the individual for the purposes of guidance the observation and interpretation of the quality of the performance is at least as important and in many cases more important than the quantitative score.* The specific abilities and the specific skills referred to in the discussion of the preceding principle are in many cases lost in the quantitative statement of accomplishment on a test. The degree to which they are to be found in the individual is revealed in the quality of his performance, which can only be seen through careful observation by a trained examiner. . . . .
>
> (9) *The validity of a test or of a battery of tests as measures of a trait or traits significant in guidance can only be determined by a correlation of test results with "an adequate criterion of ability to succeed in the work in question."*

This illustrates the logical contradictions into which the clinical point of view leads. Every ability of vocational importance must be measured by tests constructed for the purpose and yet some of these abilities cannot be measured by tests. The validity of each test as a measure of a certain ability can only be determined by correlating the test results with a criterion and yet the assumption is made without resorting to such correlation that the test score is not an index of the ability and that the extent of the ability is revealed direct to the psychologist.

THE PREDICTION OF A DISPARITY BETWEEN SCHOLARSHIP AND INTELLIGENCE

J. B. MINER

University of Kentucky

The problem of predicting special traits requires first that the special trait shall be defined if possible in objective form. The trait here studied is that of the disparity between intelligence test records and scholarship for a group of 456 college students whose scholarship records were available for the first two semesters, and who had taken the Army Alpha Test. Disparity was measured by the difference between the records of a student in terms of the standard deviation of the Alpha and of the scholarship score for each semester. So far as the trait of disparity is concerned the group was divided into those students who were higher in scholarship than the intelligence test record would anticipate, and those who were lower. The scholarship-higher and the scholarship-lower groups were then each divided as to the degree in which they showed the trait. Three degrees in each direction were defined: First, the extreme *disparity cases, those who showed consistently for two semesters* a disparity of at least one time the standard deviation, or one sigma; second, the medium cases, those who showed consistently at least 0.5 sigma but not consistently over one sigma; third, the slight disparity cases, those who showed consistently some disparity but less than 0.5 for at least one semester. The students who were inconsistent in the direction of their disparity, for the two semesters (105 in number), were classified on the basis of their average disparity for the two semesters.

An important practical problem is to find the students at the beginning of their college course who will be likely to do work far below their intellectual capacity. The studies made by Stone,[1] by Laird,[2] and by Poffenberger and Carpenter,[3] and also the present study fail to solve this practical problem. A simple way to explore the situation for leads is to use scatter diagrams in the form of correlation tables for a random group for the tests which seem best to separate the medians of the extreme groups. The different degrees of the trait may be shown by different colors of ink. Borderlines may be drawn which seem best to set off the cases which it is desired to select. Various combinations of the best tests thus found may then be tried. At first it is better to work with a small exploratory group of say 50 cases. The solution of the problem would ultimately require weighting the tests thus discovered by a partial regressive equation. Much preliminary work promises to be necessary before such refinements need be tried.

Let us first take up the results from trying to predict disparity by the total Alpha score and by the separate scores in each of the eight tests of that series. This is the only random group studied which contained the whole range of disparaties and which was large enough to try out the method thoroughly. It is essential to use a random group in order to apply the results to such a practical problem as selecting extreme cases of a trait. This should be emphasized since much time may be wasted on similar problems by attempting to discover borderlines for extreme groups by tabulating only cases showing the trait in extreme degrees. For example, if one plotted only the extreme scholarship-higher and extreme scholarship-lower groups and determined borderlines which definitely separated these groups, these borderlines would not serve to pick out the extreme cases from the rest of a random group who might have identical scores on the tests. The real problem is to find the cases of extreme disparity. When groups showing only extreme and opposite degrees of a trait are used they serve only to test whether a series of borderlines will distinguish these extreme groups from each other, not whether they will separate the extreme cases from others. The results with the Downey Will-Temperment Tests, which will be reported later, were obtained only for extreme disparity cases. They, therefore, serve only to check types of profiles which might be thought to predict disparity. On the other hand if extreme disparity cases could be found with any combination of the separate Alpha Tests the method may be trusted to have discovered such borderlines.

---

[1] Stone, C. L., Disparity between Intelligence and Scholarship, Jour. of Educ. Psychol., 1922, XIII, 241–244.

[2] Laird, Donald A., A study of Some Factors Causing a Disparity between Intelligence and Scholarship in College Students, Sch. and Soc., XIX, 290–298.

[3] Poffenberger, J. T., and Carpenter, F. L., Character Traits in School Success, Jour. of Exper. Psychol., 1924, VII, 67–74.

A word should first be said about the general relations between the Alpha Test scores and scholarship at Kentucky, since these indicate the reliability of the measures used. Credit for gathering the data and for calculating the correlations in this paper should go to E. J. Asher. His master's thesis included also systematic interviews and the individual Downey Test profiles for over fifty extreme disparity cases. For the 456 students in the College of Arts and Sciences studied in this paper the correlations between test and scholarship score were 0.44 for the first semester, 0.40 for the second semester. The reliability of the scholarship score is indicated by a correlation of 0.67 between scholarship for the first and for the second semesters. Most important in the study of disparity was the fact that disparity for the first and second semesters correlated 0.45, so that the records were undoubtedly getting some sort of complex trait which did not disappear after a semester in college. This was true for the group as a whole even when the records had not been corrected by a second intelligence test. Such verification of disparity is undoubtedly important since it was found that about one in five cases dropped out from the extreme disparity groups where the disparity was checked by a second intelligence test. The absence of verification by a second test record accounts easily for the reliability coefficient for disparity not being over 0.45 with this larger group.

The first suggestion which was tried out was that the variability of the students from intelligence test to scholarship might be an indication of a fundamental variability which would show itself also in the variability of the same student from test to test in the Alpha series. This was examined by taking the inter-quartile difference in sigma scores for the eight Alpha Tests as a measure of each student's variability and correlating this with the disparity between total test score and scholarship. Individual variability on the eight tests of the Alpha series proved to be no measure at all of disparity. The correlation with disparity for the first semester was −0.05, and for the second semester −0.02.

Taking up next the prediction of the direction of disparity by means of scatter diagrams, from a sample group of fifty cases, the two best tests for discriminating the direction of disparity were found to be Test 5, Mixed Up Sentences, and Test 8, Information. Regarding as favorable for scholarship relative to tested intelligence a score below 18 on Test 5, and below 24 on Test 8, and as unfavorable a score of 18 or above on Test 5, it was possible to properly place, as to the *direction of disparity*, 30 out of 35 cases which fell outside these limits. This percentage of success was thus 81 per cent. When these criteria were checked for the 289 cases outside these borders among the 451 cases, the number of correct predictions was found to be 78 per cent. The chances were thus 4 to 1 that the direction of disparity could be predicted for cases falling beyond certain borderlines. Supplementary records for the other Alpha Tests and the total Alpha scores were hardly worth while for the prediction of the direction of disparity.

It is a much more difficult problem to pick the cases that may be expected to show an extreme disparity of at least one sigma for two semesters. This is, however, the main practical problem, since a less disparity cannot be assumed to indicate a student who should be dealt with individually. The attempt to find extreme cases was made only for the scholarship-lower group. Here the best criterion proved to be a score of plus one or more sigma for the total Alpha score as likely to go with extreme scholarship-lower. There were 72 cases beyond this borderline among the 456. Seventy-five per cent of those 72 cases would be rightly predicted as to the direction of disparity, 55 per cent as showing at least medium disparity (0.5 sigma or more); but only 32 per cent of them were of the extreme degree of disparity sought, that is, of the type for whom special training would be clearly worth while. It is also to be noted that this borderline failed to find 17 out of 40 cases with extreme-scholarship-lower even if we might deal with those which it did find. When the above borderline for the total Alpha score is supplemented by the two best individual tests for distinguishing scholarship-lower, Tests 5 and 8, the situation is only slightly improved. Reducing the 72 cases to 31 who are also above scores of 20 on Test 5 and above 26 on Test 8, we find that the chances are 23 to 8 that such a case will show at least medium scholarship-lower, a disparity of 0.5 sigma or more. For these 31 cases, however, the chances are 19 to 12 that they will not show a disparity of 1 sigma. Under the most favorable determination of borderlines, with the chances still approximately 2 to 1 against a case being of the extreme disparity sought for special training, it is clear that other types of tests than those in Alpha should be tried.

Since the Downey Will-Temperament Tests apparently afforded one of the best instruments yet devised for getting objective records of a variety of personal traits, it was decided to try them with very carefully selected groups of extreme disparity. To be a member of these groups the student must have been tested by a second intelligence test; second, have consistently maintained a disparity of at least one sigma for each semester; and, third, be available for an interview and an individual test with the Downey series. Only 32 cases were found which met these requirements. They are much more rigid than have been applied by others who have studied the problem.

By inspection it did not seem possible to distinguish any typical differences in the general profiles even for these extremely different groups, 16 with extreme and consistent scholarship-lower, and 16 with corresponding scholarship-higher than tested intelligence. However, it might be that Dr.

Downey, who has worked so long with such profiles might find patterns in them which escaped other observers. She very generously consented to attempt to place these 32 cases in two groups: Those from whom she would expect scholarship lower than tested intelligence and those from whom she would expect scholarship higher. In returning her conclusions she qualified them by the statement that her judgment was based largely upon unpublished work which she had begun on the problem but in which she had used the group test rather than the individual test. She did not know whether such results would carry over to the individual Will-Temperament Test. She also supplemented her judgments by qualifications as to whether the student had a high or low record in tested general intelligence. It is obvious that the Downey Tests were not devised with such a problem in mind, so that it would be wholly unfair to regard this trial as a test of the Will Temperament diagnosis for other purposes or of the Downey group test. Too much credit can hardly be given Dr. Downey for her pioneer work in this fascinating field.

Considering Dr. Downey's conclusions as to which of these two extremely opposite groups a case belonged, both with and without her allowances for intelligence test records, her guesses proved to be right in just half the cases, just what a random sorting of the cases would have given. This may mean only that criteria from the group Will-Temperament Test should not be transferred to the individual Will-Temperament Test for this problem.

Knowing in which group each of these 32 extreme disparity cases belongs it is a comparatively easy matter to discover borderlines on scatter diagrams which will separate them all properly. The simplest criterion was found to be the total Alpha score, the extreme scholarship-higher students in this group with one exception were all below a score of 126 on Alpha. With a single exception also, those with scholarship-lower scored above 141 on Alpha. The two exceptions, between 126 and 140 on Army Alpha may be properly classified by being above and below the median score of 5 on the Downey Speed of

Movement Test. The two Downey Tests which seem best to discriminate these extreme disparity cases are Speed of Decision and Motor Inhibition, but combined they fail to reach 7 of the 32 cases.

The important conclusion to emphasize is that test records on extremely different groups are of little significance until comparable records are available for a random group of students. The above tests and borderlines must not be used to predict disparity since many other students with the same scores would show little disparity or an opposite tendency. Work with extreme cases alone can never solve the problem of predicting extreme cases of any type. This is demonstrated by the fact that the Alpha total score would easily separate 30 out of the 32 extreme disparity cases; but if these borderlines were used to predict extreme disparity, they would be wrong for four out of five cases in our large random freshmen group beyond the same borderlines. The borderlines which successfully separated the extreme cases are *apparently those which distinguish very good test records from very poor records, and therefore leave the biggest opportunity for a disparity from scholarship.*

It is of interest to note that the interviews by Mr. Asher were directed to finding, if possible, what seemed to be the main reasons for the extreme disparity. The results, as subjectively evaluated by him, indicated, for the 27 extreme scholarship-lower students on the basis of their first semester scholarship, that the primary influence was apparently not environmental. For 11 cases it was lack of interest in the college work, for 6 slow adjustment to the new environment, for 5 other character defects of various kinds, and for only 5 was it unusual outside influences or illness. For extreme and persistent cases, it would, therefore, seem that scholarship-lower is based on combinations of personal traits which should be discoverable by tests.

A STUDY OF THE COMPARATIVE VALUE OF NINE PERFORMANCE TESTS

KATHARINE MURDOCH

Institute of Educational Research, Teachers College, Columbia University

Performance tests, from the Pintner Paterson series, are relied on very largely in determining the intelligence of non-English-speaking persons. Frequently, due to lack of time or material, a few only of such tests are relied upon.

Does it make a great deal of difference which ones are used? Do some have a constant tendency to yield mental ages higher or lower than the median mental age which all of them if used, would yield? Are some more variable than others in comparison with such median mental ages?

This study helps partially to answer these questions for nine of these tests. The results from them are compared not with the median mental ages which the use of all the series would produce, but with the medians obtained from the use of only these nine tests taken together.

The nine tests are: Mare and Foal, Seguin, Five Figure Board, Two Figure Board, Casuist Form Board, Manikin Test, Feature Profile Test, Picture Completion and Cube Test. Time, not errors, was scored for those tests where scores can be arrived at in either way. The test materials were obtained from Stoelting and Company.

These tests were all given to a group of 29 young unmarried mothers at the San Antonio Mission Home and Training School by the writer and one other psychologist and an assistant, in the course of the Texas Mental Hygiene Survey which was conducted by the National Committee for Mental Hygiene in the fall of 1924. The method of giving and scoring the tests described in "A Scale of Performance Tests" by Pintner and Paterson, was followed. Half of the tests were in charge of one psychologist, and half in charge of the other. Half of the subjects were examined first by one of these psychologists and half by the other. One of the psychologists rotated the order of giving the tests she used; the other one did not. The Two Figure Form Board thus came to be given first of all the tests, to half of the subjects and the Casuist Form Board second. This latter and the Manikin, Feature Profile, and Cube Tests never were given first. Each of the other tests was given first to about one-eighth of the subjects.

The median mental age from all the nine tests was calculated for each subject and the deviation from this of the mental age determination of each test. The average of these deviations taken algebraically appears in the first row of the accompanying table. This shows, for each of these nine tests, how far it tends to over- or under-estimate the mental age, as this is determined by all of them. Thus we see that the Mare and Foal Test tends to give a mental age 1.89 years too high, the Seguin Board, one 2.69 years too low, the Five Figure Board, one only +0.45 of a year away from the median, etc. Of the nine tests studied, the Five Figure and Two Figure Form Boards and the Picture Completion Test yield the best results for subjects of this mental age (average for the group, 11.7 years, σ2.31). Mare and Foal, Casuist Form Board and Feature Profile give over age results; Seguin Form Board, Manikin, and Cube Test all decidedly underrate the subjects.

TABLE 1

	MARE AND FOAL TEST	SEGUIN	FIVE FIGURE BOARD	TWO FIGURE BOARD	CASUIST FORM BOARD	MANIKIN TEST	FEATURE PROFILE TEST	PICTURE COMPLETION	CUBE TEST
Average deviation from median M.A.	+1.89	−2.69	+0.45	+0.05	+1.41	−4.90	+1.31	−0.41	−2.27
Average amount of deviation, up and down from median M.A.	2.31	2.83	1.41	1.72	2.31	4.00	2.14	1.45	2.62

The reliability of the results just quoted is measured by the coefficient 0.89. This is the correlation obtained for the nine tests between the sum of the scores of 15 of the subjects and the scores of the other 14, by the method of difference in relative position.

The relative variability of scores by the nine tests from the median mental age, is shown in the second row of table 1. It is the average for each test of the amount of its deviations from the median mental ages, taken arithmetically. It shows that the Mare and Foal Test varies on the average up or down, 2.31 years from the median mental age; the Seguin Test varies 2.83 years; the Five Figure Board varies 1.41 years, etc. The reliability of these results is not very great, the coefficient of correlation between the results on fifteen subjects and on the other fourteen is 0.42 only. They indicate, however, that the three tests, Five and Two Figure Board and Picture Completion, which are best of the nine because they average nearer the median mental age, are best also because they show less variability in any direction from the median than do the other six tests.

MINOR STUDIES IN THE PSYCHOLOGY OF ADVERTISING FROM THE LABORATORY OF INDIANA UNIVERSITY

XIII. Pictures of People in Magazine Advertising

HARRY D. KITSON AND IRVING ALLEN

Indiana University

Much is said in the literature on the technique of advertising about the value of injecting "human interest" into advertising copy. One way of doing this which is recommended is to use pictures of people. An attempt to measure the value of these was made by Nixon[1] who, working under laboratory conditions, found that his subjects looked at advertisements containing people more often than they looked at advertisements containing pictures of objects, and that they gazed at the former for longer periods of time; in one series of experiments 11 per cent longer, on the average, and in another series, 25 per cent.

In order to find out to what extent advertisers are using and have been using this device, investigation by the historical method is required. No such investigation is on record, though the casual observation is made by Nixon that "a survey of current magazines shows that over 65 per cent of full page magazine advertisements contain pictures of humans as a prominent feature, and it is probable that this use is on the increase." In order to gather exact information on this point the writers examined 500 consecutive illustrated advertisements (all sizes) in each of the four periodicals of the issues shown below:

---

[1] Nixon, H. K., Attention and Interest in Advertising, Archives of Psychology, Columbia University, New York, 1924.

Nixon, H. K., Two Studies of Attention to Advertisements, Jour. of Applied Psychol., June, 1925, IX, 176-187.

Saturday Evening Post, April 18, and May 2, 1925.
Literary Digest, March 14 and 21; April 4, 11 and 18; May 2 and 9, 1925.
Womans Home Companion, May and June, 1924; January and May, 1925.

In order to discover the trend over a lapse of time, 500 consecutive illustrated advertisements were examined in past issues of the *Literary Digest* as follows:

From July 3 to September 4, 1915
From July 7 to October 13, 1906

TABLE 1

Frequency with which humans appear in illustrated advertisements

MAGAZINE	TOTAL NUMBER ILLUSTRATED ADS EXAMINED	NUMBER SHOWING MALES	NUMBER SHOWING FEMALES	TOTAL NUMBER SHOWING HUMANS	PER CENT SHOWING HUMANS	PER CENT (OF LATTER) SHOWING MALES	PER CENT SHOWING FEMALES
Saturday Evening Post........	532	300	216	384	72.2	70.7	56.2
Woman's Home Companion.....	500	141	328	373	74.6	37.8	87.6
Literary Digest................	513	249	187	322	62.8	77.3	58.0
Literary Digest, 1915...........	500	237	108	274	54.8	86.5	39.4
Literary Digest, 1906...........	500	152	94	219	43.8	69.4	42.9

Table 1 shows a frequency of pictures of human beings in the illustrated advertisements examined as represented by the following per cents: *Saturday Evening Post*, 72; *Literary Digest*, 62; *Woman Home Companion* 74. The average frequency of all is 69.

The data were so tabulated as to show the number of advertisements containing pictures of males and the number containing pictures of females. (Some contained both.) The per cent showing males exceeds the per cent showing females in both the *Saturday Evening Post* and the *Literary Digest*, the figures being 79 per cent males and 56 per cent females in the former, and 77 per cent males and 58 per cent females in the latter. This may mean that the majority of the readers of these periodicals are men; or it may mean that the articles advertised in these periodicals are predominantly those used by men.

In the woman's magazine, as would be expected, the number of advertisements containing pictures of women is greater, the figures being: females 87, males 37. The trend of practise during the past twenty years is indicated by figures gathered from a study of the files of the *Literary Digest* for 1906 and 1915. These figures show that the use of pictures of humans has been increasing during the past twenty years. In 1906 the per cent was 43; in 1915, 54; and 1925, 62; an increase during the past twenty years amounting to about 40 per cent.

It is interesting to note in this connection that the use of pictures of women has increased more rapidly than the use of pictures of men. In 1906 the number showing females was 42, while in 1925 it reached 58, an increase amounting to 38 per cent. In 1906 the number showing males was 69, in 1925, 77, an increase amounting to only 11 per cent. This increasing popularity of pictures of the female may indicate several things: That the proportion of women readers of the *Literary Digest* has increased; that an increasing number of commodities used by women has come to be advertised in this periodical; that the number of women buyers (wage-earners?) has increased relatively faster than the number of men buyers; that it is coming to be recognized by advertisers that women do most of the buying. One might venture the hypothesis also that perhaps advertisers have been trying to use the much flaunted "sex-appeal" among the male readers of this masculine (?) periodical. To what extent each of these factors contributes to the increase in the use of pictures of females can only be a matter of conjecture.

The tendency toward use of pictures of humans in advertising copy should be viewed in connection with another fact. It has already been discovered by one of the writers that there is a growing tendency in advertising copy to describe an article in terms of its use.[2] In carrying this out the most natural medium to employ is a picture of a human being using the article.

Another factor may be simply the well-known increase that has occurred in the absolute number of illustrations. Figures already published by one of the writers have shown that the frequency of illustrations rose from 70 in 100 to 90 in 100 during the period 1907 to 1919.[3] With this increase in number together with the great advances made in the technique of illustration, would naturally go a wider use of human materials.

SUMMARY

This investigation has shown that in the representative magazines examined the number of illustrated advertisements *containing pictures of humans is on the average 69.* The figures show that the tendency in this direction has been increasing during the past twenty years, the increase amounting to about 44 per cent. In the woman's magazine examined the advertisements containing pictures of women predominated; in the two general magazines, the number containing pictures of men predominated. Pictures of women have been growing most rapidly in popularity.

---

[2] Kitson, H. D., The Mind of the Buyer, New York, Macmillan, 1921, 83 ff.

[3] Kitson, H. D., Amount and Rate of Increase in the Use of Illustrations, Journal of Applied Psychology, March, 1921, V, 12–13.

A SPEECH CLINIC CASE WITH MISCONDUCT AS A BY-PRODUCT

HERMAN H. YOUNG

Psychological Clinic, Indiana University

Wilma was brought to the clinical psychology class for examination because of her inability to speak properly and because she was a disciplinary problem both at home and at school. The arrangements for her examination were made by a student of the class who had been her teacher in the 1A grade. Although Wilma was a girl seven years and three months old and in the 2B grade of the Indianapolis Public Schools her teacher doubted if it would be of much value to the class to be present during her examination because her speech was so poor no one could understand more than a very small percentage of all the words she spoke. As it is the purpose and obligation of clinical psychology to study all types of cases, the teacher was advised to bring the girl.

The teacher reported that she had promoted Wilma to the second grade because she felt she could do the work. Her written work and comprehension were very good. She was liked by everyone. She had been put into an atypical class for reading each day on account of her speech. She hated this period. The children in their play seemed to understand her. She liked to play games out of doors. In the school room she had a horrid disposition, was a vicious child seemed possessed to pinch and kick somebody. She pouted, was not happy and was never amused at anything that ever happened in the school room.

Wilma was brought to the clinic by an aunt who reported that when Wilma was three years old while riding in an automobile with her family a train hit the auto and killed all other members of her immediate family. She was thrown a dis-

THE JOURNAL OF APPLIED PSYCHOLOGY, VOL. IX, NO. 4.

tance of about 100 feet and was considerably bruised but had no serious injuries. After this accident she was taken into the home of this aunt and has been with her ever since. The girl had never been able to talk distinctly. The aunt says that she had tried every means she could think of to help the girl but had not been able to produce any effects upon her speech. Two different physicians had been consulted. One told the aunt that the girl's speech was a matter of habit, that there was absolutely nothing wrong with the child and that she would learn to speak properly when she went to school. The other said that the girl could never be expected to speak properly because for some reason the nerve of speech was undeveloped. At the time Wilma entered school the aunt explained her defective speech to the principal and teacher stating that she doubted the value of sending her to school because no one could understand her. The principal informed the aunt that they frequently get children with speech defects and that they always work with them and teach them to speak properly in addition to giving them the regular school work. After that they get along satisfactorily. Such was the encouragement given the aunt in February, 1924, when Wilma entered school. At the time she was brought for examination in February, 1925, a year later, both the principal and the first grade teachers reported that they could not teach the child to speak properly. The principal said this is the first child upon whom they have failed entirely. Although Wilma was promoted to the second grade the general impression seemed to be that her speech was very little if any better than when she entered the first grade.

The aunt hesitated to bring the girl to the clinic because she herself had tried various ways and means of teaching the girl and had received no satisfaction, and because even after the encouragement given her by the school principal, Wilma made so little improvement that the aunt was inclined to feel nothing could be done for her speech. Some people had told her she should not pay any attention to it and that the child would outgrow it. The aunt felt she had waited long enough. She had about concluded that the child could not be instructed and that she would not outgrow it. She said, "Not one single word of her reading could be understood." Wilma could not say many of the ordinary things around home without much explanation. By way of illustration the aunt told of one occasion upon which Wilma wanted to tell her something was "yellow." No one at her home could understand her. Finally she told them it is like the sun. In all, it required half an hour for her to succeed in having the family understand that she meant to say "yellow." She pronounced her own name and address so imperfectly that had she gotten lost from home it would have been impossible for her to make a stranger understand who she is or where she lives.

In addition to the girl's defective speech the aunt was very much worried because of her behavior. The girl's general attitude and conduct were a source of as much worry as that of her speech. In fact, it seemed to impress people as the more serious of the two.

PSYCHOLOGICAL EXAMINATION

Wilma first met the examiner in his class in clinical psychology. The entire examination was conducted in front of the class. She was given the Stanford Revision of the Binet test together with performance tests to determine her general mental level and to get an index of her powers of observation, attention, imageability, attention span, and ability to learn. At the conclusion of this examination she was diagnosed as of slightly above average normal mental ability. Her mental age on the Binet was seven years and six months or three months above her chronological age. In making this diagnosis considerable weight was given to the quality of her performances on the performance tests. This was necessary because her answers to some of the Binet tests were difficult to understand while in other cases it was impossible to decide definitely what she was attempting to say. At times it required the best efforts of the class, the aunt, and the examiner to interpret what she might have meant by the sounds she made. Often she was asked to repeat her statements three and four times before any one of those present felt safe to venture a guess.

After the above tests were completed the examiner had Wilma say the sounds: "ah," "ee," "oh," "oo," and that of the pronounciation of several simple words. This constituted the entire examination.

During the entire period of the above examination, Wilma's speech was studied very carefully by the examiner although unknown to Wilma or the class. This study was directed to her speech in an effort to determine whether there might be any physical condition responsible for her speech. Such physical condition might consist of malformations of such a nature that it would be impossible to utter the sounds essential in speaking English. Or they might consist of paralyses which made it impossible for Wilma to control the organs of speech in attempting to pronounce words. Keeping these possibilities in mind the examiner noted that Wilma apparently was able to control the sounds essential to the speaking of good English but did not have these properly coördinated so that the language she spoke was English. This convinced the examiner (1) that there was no serious physical abnormality present, (2) that her unintelligible speech was the result of improper speech habits and (3) that if the known laws of learning and habit formation were intelligently applied in her case she could easily learn to speak correct English. For this reason the examiner predicted that if Wilma were taught with these objectives in mind and consistently employed that a 100 per cent cure could be effected within a year or a year and a half. When the examiner informed the aunt of his conclusions she did not appear very much impressed by them and gave the attitude that she had heard such encouragements previously and that the examiner would later learn of his erroneous conclusions. She was, however, willing that arrangements be made for the girl to be taught under the examiner's direction.

The success in a case like this depends upon the ability of the instructor to make a correct diagnosis of the condition and then to direct the overcoming of a wrong set of habits and the building up of a correct set by direct and easy guidance of the learning processes. To do this a new diagnosis must be made at each stage of the successive changes through which the individual being taught must go in making the transition from the old to the new habits. This means that no cut and dried method is applicable and that the effectiveness with which success is achieved will depend upon the teacher's ability to make accurate detailed observations and correct interpretations of these observations. It also means that the teacher must have such a knowledge of how old habits are displaced and new ones formed that it functions in the directing of the teaching process as indicated by the interpretation of the observations made.

Because it was impossible to secure a teacher with training in corrective speech work it was necessary for the examiner to select an untrained person and direct the work. Miss Dorothy G. Daugherty, a graduate student in clinical psychology, was selected because of her training in clinical psychology and her acquaintance with the fundamental laws of learning and habit formation. It is interesting to note the fact that throughout her college course Miss Daugherty had intentionally avoided all education courses because she was determined never to teach. When Wilma's case was explained to her she agreed to undertake it but made it clear that she had never intended to teach and had never given any attention to teaching methods and practice.

After Miss Daugherty had agreed to teach Wilma and Wilma's aunt had also agreed to the teaching, the aunt seemed to feel that even though success was unlikely Wilma's speech would be given due attention. According to the plan Wilma got from one-half to three-quarters of an hour instruction every Saturday morning from February to June, 1925. Lest the examiner had forgotten about Wilma's conduct, the aunt asked about suggestions for her discipline. The examiner advised her in as strong terms as possible to "forget it," and then explained that the speech was the primary condition to be considered and that Wilma's conduct might cease to be a problem providing her speech is corrected.

The aunt was naturally concerned about Wilma's behavior because many people blamed her for not disciplining her properly, while other people felt that Wilma was an abnormal child. The aunt reported that various forms of punishment had been tried, but in no case was Wilma whipped. Wilma was in trouble in school so frequently that reports of her bad conduct were sent home on an average of about once a week. At times during school hours she would just make loud, ugly noises.

The week after the clinical examination and diagnosis were made both Wilma and her new teacher in speech appeared for their first lesson. One-half hour of time was spent by the examiner on giving Wilma her first speech lesson and instructing her teacher how to continue the lessons. As different aspects of the problem were revealed the examiner explained what he considered to be the fundamental principles to both pupil and teacher. After this first lesson the examiner did not spend more than fifteen or twenty minutes with Wilma. During this first lesson at least two significant features of the case were brought out. It was found that "chair" was one of the words Wilma could not pronounce intelligibly. It was further found that she could with very little practice give the "ch" sound and the "air" sound, each independent of the other but could not possibly combine them into the word "chair." She was therefore induced to pronounce both these separate sounds as independent words, decreasing the time elapsing between the pronunciation of the two parts of the word. She was, however, never able to pronounce the word during this first half hour. Whenever she was asked to say "chair" she invariably pronounced it by her old method. This is illustrative of her difficulties. During this first lesson she was asked for a word which she could not pronounce. In reply she tried to pronounce a phrase which she pronounced so poorly that her aunt could understand it only after several repetitions. "Thank you" was the phrase she was attempting to say and was very anxious to learn correctly. During this first half hour we were unable to teach her to say "thank" but the examiner carefully told her how to place her tongue and other speech organs. Then by way of a side remark the examiner told the aunt that the best thing for Wilma to do would be to practice that during the week and if possible to practice it in front of a mirror where she might observe her own movements. That this method of trying to spur Wilma on succeeded was evidenced a week later when she appeared for her lesson and could say "thank you" very well.

Wilma took hold of the work so actively that the examiner realized she would make very rapid progress, in fact would probably make much more rapid progress than would ordinarily be expected. For this reason a sample of her reading was taken upon her fourth visit to the Clinic. By this time she had already made sufficient improvement to pronounce correctly some of the words which she had been unable to pronounce previously, that the sample of her reading taken at this time is not as poor as it would have been had it been taken upon her first visit to the Clinic. The aunt stated that before Wilma came to the Clinic no one could understand a single word of anything she read aloud. A copy of the passages to be read and an attempt at a spelling of the words as they sounded to a listener when she pronounced them are quoted below one above the other. Even though her pronunciations were incorrect she was able to understand and interpret everything she read.

Original: "A long time ago a dog had the right to eat
Wilma's reading: "A om ti ago a dodge ad de wite to eat

all the meat that fell on the floor. That was a fine time for the dog.
all de meat at fall on the for. That wa a fi time far de dodge.

He was fat and happy.
E wa fa an appy.

"He was afraid someone would take this right from him. So he
" E wa alade sahvon vood take thit wite from cem. Szo e

asked the king to give him a paper and the paper said that the
uk de kid to gi cem a baber an de baber sai at de

dog could have the meat. The cat said she would keep the paper
dodge coo ave de meat. The ca sai se vood teep de baber

for the dog."
far do dodge."

Five in One Pod

"There were once 5 peas in one pod. They were green and the
"There war once 5 bae in hon bod. Zo var feen and de

pod was green. So they thought all the world was green. Why
bod vor feen. Zo de taut off de worl wa feen. Vy

shouldn't they have thought so? The pod grew and grew as they sat
sooda ay ave taut so? The bod brew and brew as ze gat

there in a row."
there in a row."

On the same day that this sample of reading was secured Wilma was asked to go before the class in Clinical Psychology again and read for them. At first she refused to go. She went only after considerable urging in which it was explained to her that the purpose was for her to let them hear how she reads now and then later when she can read properly she could go before them again to show them how much she had improved. Three months later on the next to her last trip to the Clinic she asked permission from Miss Daugherty to go in and read before the class. This is an indication of the change in the girl's attitude toward life in general. Another indication of this change in attitude is the fact that the school did not report her to her aunt for bad conduct during the last two months of the school year. Prior to that a report for bad conduct had been sent to her home on an average of about once a week. The aunt estimated that at least 75 per cent of the causes of worry about Wilma's bad conduct had disappeared.

The teacher reported she was much happier in school and that now she wanted to talk to the teacher and help her. She never wanted to talk or help before. When asked why she behaves so much better now than she used to, she said, "Guess 'cause I wanted to be naughty," but she could not tell why she wanted to be naughty before and does not want to be naughty now.

The great improvement which Wilma made in her speech during the four months from February to June is indicated (1) by the aunt's statement, that "There isn't anybody who has any trouble understanding her now," and (2) that whereas in the two passages of 118 words quoted above she pronounced only 30 words or 25 per cent intelligibly late in February, she pronounced 111 words or 93 per cent intelligibly early in June. The only words she mispronounced entirely on this latter date were "had," "fell," "his," and "all." She had considerable difficulty with "right," "world," and "why." Her improvement in conversation was greater than that in reading aloud.

It is not known exactly why Wilma did not speak correct English. Indications are that she developed an individual pronunciation of English meaning words through contact with Polish servants in her home when she was a child. Her pronunciations, however, did not impress one as that of a foreigner trying to learn English. They seemed more like that of an individual language. From that angle it hardly seems fair to the facts to speak of her as a speech defective, unless we wish to classify all users of other languages as speech defectives.

Her rapid acquiring of the English language was due to several facts.

1. That she had English meanings and comprehension, but imperfect English pronunciations.
2. That her condition was correctly diagnosed at the beginning and correctly re-diagnosed at every step of instruction.
3. That she was therefore instructed from the beginning according to known laws of learning and of habit formation with the instruction always aimed directly at the needs indicated by the re-diagnoses of her progressively improving English speech.

4. That she was of normal mental ability somewhat above average.

5. That she was so conscious and painfully aware of her unintelligible speech that she came with an intense desire to learn to speak correctly.

6. That this desire to learn persisted throughout the learning period because for the first time in her life her efforts to learn to speak correctly were being crowned with real success.

7. That her aunt who stayed with her during every period of instruction learned enough about Wilma's difficulties and the methods of overcoming them to give intelligent and effective help between periods of regular instruction at the Clinic.

This report of the work and accomplishments with Wilma is published because there are many other children who are victims not only of detrimental speech habits but of other equally serious habits which will respond equally as satisfactorily to proper instruction if diagnosed and treated early enough. Early correct diagnosis and proper treatment of real causes can do wonders. Failure results from incorrect diagnoses and attempts to treat symptoms as if they were causes. This latter is what would have happened had Wilma been given "strict discipline" for her "bad" conduct, instead of being instructed in correct speech habits. "Bad" conduct in children is always a symptom which can be removed by discovering its underlying causes and treating them. In this case the cause was improper speech habits. In other cases it is malnutrition, or defective vision, or defective hearing, or mistreatment at home or in school, or feeblemindedness, or insanity, or indigestion, or being petted too much, or physical deformity, or any one of a thousand other things. The variety and inter-relation of possible causes is unlimited. The variety and adaptations of methods of treatment must be just as large and unlimited as that of the causes. The particular combination of causes operating in each individual case must determine the method or methods of procedure to be employed. It is easy to understand how in Wilma's case her unintelligible speech was directly the cause of her misconduct. People could not understand her. She was asked to repeat her statements over and over and then people with an expression which betrayed their attitude toward her pretended to understand. Could a child who had to spend half an hour to make her family understand she meant to say a thing was "yellow," reasonably be expected to be happy, sweet-dispositioned and invariably obedient? Suppose she had been severely "disciplined," instead of instructed to speak correctly, what could have been expected of her as an adolescent and as an adult?

THE INTELLIGENCE OF FULL BLOOD INDIANS

THOMAS R. GARTH

ASSISTED BY T. J. SERAFINI AND DEWEY DUTTON

University of Denver

The American Indian as a subject of study for several years has attracted the attention of students of racial psychology. However, there have been handicaps to satisfactory investigations.

For one thing it has been rather difficult to obtain subjects of pure Indian blood in sufficiently large numbers to make the results very significant. Another difficulty has been in finding for the results so far obtained on these full blood Indians a basis of comparison with white subjects measured by the same tests.

The latter difficulty is largely, though not altogether, due to the same fact, insufficiently large groups. Norms for whites on various standard tests are too low on the age scale to be of much assistance because of the advanced age of the Indians so far tested. That is to say, the Indians are too old to be compared with the whites. Very few full blood Indians of years eight, nine and ten have been tested. They are in fact found to be around twelve, thirteen and fourteen years old when they are obtained in numbers of any magnitude in testing expeditions and we have rather unsatisfactory white norms for these ages. Sometimes the Indian ages run as high as eighteen and nineteen years and occasionally twenty-two and twenty-three years. What is more, an adequate number of the Indians have not, age for age, the educational equivalent of the white subjects. If we take Indians of the fourth grade they will be found to be about fourteen years old on the average and the whites of the same grade will be around nine years. If we undertake then to control age the factor of education "runs at large" and if we try to control the educational factor the age element in the problem "runs riot." What is needed is some means for controlling the one or the other. In the light of previous practice in general and for the sake of convenience here the practical thing to do in the study of Indian intelligence is to control the age factor if this can be done. This is now made possible by considering, if only tentatively, the mental age equivalents of group intelligence test norms for whites as obtained from large numbers in the United States and determining what the intelligence of an Indian would be under the same circumstances. This, to be sure, disregards the factor of social status which, as between the average white in the United States and the average Indian is certainly different, qualitatively and quantitatively. This is true of Indian children in the United States Indian Schools as well as in their homes on the reservations, but one would be inclined to think it less so in the former case. One has but to visit the Indian in his hogan, his tepee, or pueblo hut or to see him in his community life to be convinced of the meagreness of these as any sort of adequate preparation for the ways and customs of civilization. It has been said to the writer by an Indian school man and reservation administrator of much experience that he would not, if it were in his power, destroy the hogan with its home atmosphere and teaching since from it, he confidently believes, have come some of the "finest characters" he has known. It is true one has to admire the Navajo for his rugged honesty and thriftiness, the Ute for his straightforwardness, the Pueblo for his amiability, and the Sioux and Arapahoe for their courage and endurance. But these, while they may be provided for severally, in the hogan, the adobe hut, the tepee and in the community life, possibly through both heredity and social environment, are not necessarily the traits measured by intelligence tests.

Some of the most commonplace experiences for the average white are lacking to the average Indians. For instance the hogan and tepee have no "chairs" "tables," "cellars," "bureaus" "teddy-bears," "flower-pots," "cooking stoves," "books," or "clocks." Once in a while we find an iron bed, a trunk, a cooking stove, in a Pueblo home. On one such occasion we marveled at the presence of a kitchen cabinet. We will say nothing of combs, pianos, electric lights, looking glasses and lanterns, such common appointments of the home of the white child. As we approach nearer to the white settlements we find them and that the Indian child in a United States Indian School has some familiarity with such as these is a fact. But a knowledge about such commonplace things and their significance in civilization is presupposed in those who are to pass satisfactorily the white man's intelligence tests.

In the present study we have endeavored to get round all the difficulties mentioned and have been successful in meeting all of them but the last one—that of social status. We have met this as best we could by taking children from a white man's environment as found in the United States Indian School.

With respect to the factor of selection we cannot say that ours is strictly a random sampling. The large number of cases in our study selected from a wide geographical area would indicate a random sampling. Still we have to admit that the Indian boarding school would be in some instances an instrument for selecting the brighter Indians. This would obtain in the school far removed from the Indian home. Nevertheless many schools are located immediately on the reservation and here the factor of selection would not really be selective any more than it is in many white rural communities. The subjects of our study are, we think, fairly representative of their Indian communities, since all the United States Indian Schools are fairly close to the communities from which our subjects come.

In this study we have several objectives: (1) To ascertain the approximate intelligence quotient (IQ) of the full blood Indians in the United States. (2) to see if the IQ differs from grade to grade; (3) to find what correlation exists between this approximate IQ and formal education; (4) to ascertain the median mental age for each school grade of these Indians.

The total number of subjects for whom we have obtained IQ's and mental ages in this study is 1102. Fifty-two of them are from the ninth grade and may not be included legitimately in the large group from which we obtain a tentative norm. We have used these fifty-two ninth graders merely to see if their superior education will favor them above the others in the grades below them. If we deduct the number of the ninth grade group from the total number tested we have a group of one thousand and fifty full blood Indian children of plains and southeastern tribes and those of the southwestern plateau tribes as the Pueblo, Navajo and Apache tribes.[1] The Indian schools visited were the United States Schools located respectively at Chilocco in Oklahoma and Albuquerque, Sante Fe and Shiprock in New Mexico.

The facts about degree of Indian blood and age are not as securely established as one might desire but since such facts are given and accepted by the United States Government we feel that on the whole they are correct. At least we find no other source of information than that of which the government avails itself. The tests used were the National Intelligence Tests, Scale A, Form I.

Table 1 gives for these full blood Indians the distributions of the approximate IQ's of the fourth up through the eighth grade; the sub-totals of these; the totals for the one thousand and fifty; and the IQs also for purposes of comparison of the fifty-two ninth graders. We have given here also the medians and Q's for the grade groups and the larger total group.

An examination of the total distribution in table 1 reveals that we have an approximately normal curve but for the high frequency at the lower end of the scale. This situation may be explained rather satisfactorily by noting that the high spur is due to undistributed IQ's. There were so many Indian children scoring lower than any authoritative mental age supplied below that of eight and one-half years. This would seem an unusual situation. In the fourth grade there were one hundred and thirty-eight Indians with mental ages below eight and a half years. In the fifth grade there were thirty-six, in the sixth grade twenty-two and in the seventh

---

[1] Part of the data for the group of one thousand and fifty was obtained on an expedition financed by the Grants Committee of the American Association for the Advancement of Science in the Spring of 1921. The rest of the data was obtained on subsequent expeditions.

TABLE I

Distribution of N. I. T. approximate IQ's for separate grades and total for 1050 full blood Indians

IQ (APPROXIMATE)	GRADES						
	Fourth	Fifth	Sixth	Seventh	Eighth	Total	Ninth
135–139			1			1	
130–134							
125–129	1					1	
120–124	1					1	
115–119							
110–114					2	2	
105–109		2	2			4	
100–104	2	1	2	2	1	8	1
95– 99	1	3	9	3	6	22	2
90– 94	3	9	4	8	4	28	2
85– 89	6	7	14	13	12	52	11
80– 84	18	19	21	30	6	94	12
75– 79	18	20	39	36	9	122	14
70– 74	26	28	59	25	9	147	6
65– 69	23	40	58	29	9	159	4
60– 64	17	32	37	17	4	107	
55– 59	15	28	24	9	0	76	
50– 54	6	16	5	1		28	
(Below or) 0– 49	138	36	22	2		198	
Totals................	275	241	297	175	62	1050	52
Median IQ............	50.0	66.1	70.2	75.6	80.0	68.6	80.08
Q, semi inter-quartile range........	—	9.3	7.1	7.2	8.8	10.3	5.1
Median M.A...........	8.5	9.2	10.6	11.6	13.1	—	12.7
Q, semi inter-quartile range................	—	1.0	0.9	1.0	1.1	—	0.9

grade two. It would be unexampled to find such a situation among white children where half of the children of the fourth grade have a mental age of eight and a half years and less. Calculating a median IQ for such a distribution is beset with some precariousness. We have undertaken it however and find the IQ for the fourth grade to be around 48.5. The median IQ for the fifth grade jumps to 66.1, that for sixth grade is 70.2 and for the seventh and eighth grades it is respectively 75.6 and 80.0. Ninth grade students are not supposed to be tested with the National Intelligence Test. But it is interesting to note that the median IQ for fifty-two ninth grade Indians is 81.8 or about the same as that of the eighth grade group of Indians. It will be seen that the median of the one thousand and fifty IQ's is 68.6 with a Q of 10.3. Calculating averages, average deviations and probable errors directly seemed an unprofitable procedure because of the undistributed IQ's and for that reason we have computed only the medians and the Q's.

TABLE 2

The per cent of cases attaining and exceeding the median IQ of the preceding grade

	PER CENT
Fifth grade on fourth grade median	83
Sixth grade on fifth grade median	69
Seventh grade on sixth grade median	67
Eighth grade on seventh grade median	64
Ninth grade on eighth grade median	51

It will be seen on examining table 1 again that there is a steady rise in the median IQ in going from grade to grade, and upon availing ourselves of the facts of table 2 we see that the differences from grade to grade tend at first to be marked with the tendency becoming regularly less strong as we proceed from grade to grade. The overlapping reads 83, 69, 67, 64, and 51 per cent respectively as we compare the lower grade IQ's with that of the successively higher grades.

The computation of a coefficient of correlation between IQ and grade did not appear to be a worthwhile task because of the large number of the undistributed IQ's at the lower end of the scale. Accordingly we have taken the upper four-fifths of the cases, disregarding the undistributed IQ's, and obtained a correlation of 0.25 ±0.021. This would indicate that education while a positive factor in determining the IQ is not as great as one might expect for the upper four-fifths of the cases. It might, however, signify that only the more intelligent full blood Indians feel a disposition to stay at school since the number of cases actually decreases as we go from the sixth grade on to the ninth grade.

To be sure these one thousand fifty Indians are generally retarded educationally but it seems worth while to inquire what is the corresponding mental age as revealed by the test for each grade and how far below the mental age of the whites they fall. Table 3 gives the median mental age for each grade group for full blood Indians, the averge mental age of the white children for the corresponding grades, which are used as norms, and the differences between these Indian medians and the norms for the whites.

TABLE 3

	GRADE					
	High Third	Fourth	Fifth	Sixth	Seventh	Eighth
Indian Median................	—	8.5	9.2	10.6	11.0	13.1
Q, semi inter-quartile range....	—	—	1.0	0.9	1.0	1.1
White Av. M.A. norm in years..	8.8	9.8	11.2	12.0	13.2	14.3
Difference in years.............	—	1.3	2.0	1.4	1.6	1.2

Average difference 1.5 years between white and Indian children.

It will be seen that the differences in mental age between white and Indian children vary as follows,—least in the eighth grade, 1.2 years; in the fourth grade, 1.3 years; in the sixth grade 1.4 years; seventh grade 1.6 years, and greatest in the fifth grade, 2 years. On the average the difference in mental age is 1.5 years, disregarding the chronological age of the Indians. The ratio between mental age of these Indians and of whites of grades 4 to 8 in the United States on the average is 100 to 114 or the whites are 14 per cent better than the Indians. But if on the other hand we take into consideration the chronological age up to sixteen years the ratio of IQ's is about 69 to 100.

Briefly we may conclude as follows:

1. The approximate IQ's of full blood Indians of plains and southwestern tribes in 69.
2. There is a constant tendency for IQ's as found to increase with education.
3. The correlation of distributed IQ's and school grade is a small positive one (0.25) which may signify selection of better Indians with increasing ability to pass the test with education.
4. The mental age of these Indians and whites in the United States grade for grade stands in a ratio of 100 to 114, the whites being 14 per cent better than the Indians.
5. The social status of the Indians in a United States Indian School is more nearly that of the average white than of an Indian not so favored.
6. Because of differences in social status and temperament we cannot conclude that our results are true and final measures of the intelligence of Indian children.

A NEW POINT PERFORMANCE SCALE

GRACE ARTHUR

Psychologist, Child Guidance Clinic, Amherst H. Wilder Charity, St. Paul, Minnesota

In the course of our work in a child guidance clinic, we meet a large number of children for whom it is highly desirable to obtain a performance scale rating, and a few for whom such a rating is the only obtainable index of general intelligence.

The problem of constructing a performance scale from tests available at the present time is two-fold. First, the most useful tests must be selected, and second, results from the tests selected must be combined in such a way as to yield the most useful values.

In connection with the construction of a group point scale for the measurement of intelligence, the writer had occasion to investigate the relative usefulness of the point method used in the present work, the year method of the Binet scales, and the median mental age method of computing mental age. The degree of correlation with individual test results of mental ages obtained from the same data by the three methods was in the above order, the point method showing the most agreement, the median mental age, the least. Averages of values obtained by the three methods agree closely, but in certain individual cases the median mental age behaves in erratic fashion that gives a mental age rating three or four years below that obtained by either the point method or the individual Binet.

In the selection of tests, the discriminative value formula $\left(\frac{\mathrm{Av}_2 - \mathrm{Av}_1}{\frac{\sigma_2 + \sigma_1}{2}}\right)$ used and described in "An absolute intelligence scale: A study in method," by Arthur and Woodrow (1) was applied in a modified form. As most of the norms were taken from Pintner and Paterson (2), the median and Q values were substituted for average and standard deviation. Results from applying this formula to norms available at the present time indicate the Kohs Block Design test has the highest discriminative value, for chronological ages five to fifteen inclusive, of any of the tests considered. Sylvester's norms for the Seguin form board (time) show the next highest discriminative capacity. Porteus has not included a measure of variability in the description of his maze test. On the basis of results from 343 clinic cases, it ranks third in discriminative value. Mare and Foal (time) ranks fourth, and Knox Cube test, fifth. Of the two Healy Picture Completion tests, Picture II exhibits slightly greater discriminative value, but Picture I is more useful for lower age groups, and therefore is included in the present scale. The Manikin (score) and Feature Profile (time score), regarded as two parts of a single test, are fairly useful. Combining time and error scores of the Casuist form board into a single value, increases the practical usefulness of this test. The Two Figure form board, from the standpoint of error scores, appears to be practically useless, but from the standpoint of time scores has a limited value.

The tests included in the present point performance scale therefore, are:

1. Knox Cube
2. Seguin Form Board (time)
3. Two Figure Form Board (time)
4. Casuist Form Board (time and error combined)
5. {Manikin (score)
   Feature Profile (time)
6. Mare and Foal (time)
7. Healy Picture Completion I (score)
8. Kohs Block Design (score)

In table 1 the discriminative value of each of the above tests, is given for each chronological age level from five to fifteen, inclusive. The maze is not included because of the uncertain value of the clinic data available. These discriminative values furnish the point values of the scale. In the column headed

"Scale" is given the cumulative point value corresponding to each crude score norm. These cumulative point values constitute the point scales of the separate tests. As they are all in terms of the same measure of variability, they may be considered as being made up of comparable units. The point values for the separate tests may, therefore, be totalled for each age level. The result is a combined point scale, as shown in table 2.

TABLE 1

Discriminative value of each test of the scale for chronological age levels 5 to 15, inclusive

Knox Cube Test

C.A.	CASES	SCORE	D.V.	SCALE
5 yr.	61	2		X
6	49	4	1.60	1.60
7	56	5	1.—	2.60
8	57	6	1.33	3.93
9	70	6		
10	95	6		
11	94	7	1.33	5.26
12	101	7		
13	77	7		
14	94	8	0.80	6.06
15	62			

Seguin Form Board

C.A.	CASES	TIME SCORE	D.V.	SCALE
5 yr.	80	37		X
6	170	26	2.20	2.20
7	173	23	1.00	3.20
8	206	20	1.33	4.53
9	214	18	1.00	5.53
10	221	16	1.00	6.53
11	172	15	0.56	7.09
12	141	14	0.67	7.76
13	80	12	1.33	9.09
14	80	11	0.67	9.76
15				

TABLE 1—*Continued*

Two Figure Form Board (Pintner)

C.A.	CASES	TIME SCORE	D.V.	SCALE
5 yr.	54	200		X
6	118	175	0.23	0.23
7	115	116	0.53	0.76
8	142	62	0.71	1.47
9	95	47	0.44	1.91
10	100	47		
11	123	38	0.37	2.28
12	103			
13	55	29	0.63	2.91
14	30			
15	10			

Casuist Form Board (Knox)

C.A.	CASES	TIME			ERRORS		
		Score	D.V.	Scale	Score	D.V.	Scale
5 yr.	20	DNC		X	DNC		X
6	100	300		X	30		X
7	123	154	1.18	1.18	12	1.18	1.18
8	144	106	0.71	1.89	8	0.47	1.65
9	121	93	0.25	2.14	7	0.17	1.82
10	134	78	0.33	2.47	5	0.38	2.20
11	126	68	0.35	2.82	5		
12	56	66	0.08	2.90	4	0.28	2.48
13	44				3	0.44	2.92
14	35	58	0.37	3.27	3		

Manikin (Pintner)

C.A.	CASES	SCORE	D.V.	SCALE
5 yr.	68	3		X
6	115	4	1.–	1.–
7	120	4		
8	127	5	1.33	2.33
9	74	5		
10	25	5		
11	20	5		
12	18	5		
13	8	5		
14	5	5		
15	1	5		

TABLE 1—*Continued*

Feature Profile (Knox and Kempf)

C.A.	CASES	TIME SCORE	D.V.	SCALE
5 yr.				
6	46	DNC		
7	59	DNC		
8	77	DNC		
9	81	DNC		X
10	95	240	1. —	1. —
11	86	157	0.95	1.95
12	82			
13	68	150	0.09	2.04
14	65	132	0.21	2.25
15	34			

Mare and Foal (Healy)

C.A.	CASES	TIME SCORE	D.V.	SCALE
5 yr.	20	107		X
6	67	71	1.18	1.18
7	85	62	0.49	1.67
8	98	48	0.98	2.65
9	72	41	0.74	3.39
10	77	36	0.65	4.04
11	81	34	0.28	4.32
12	70			4.32
13	42	29	0.01	5.23
14	27			
15	16	28	0.20	5.43

Picture Completion Test (Pintner and Anderson)

C.A.	CASES	TIME SCORE	D.V.	SCALE
5 yr.	20	80		X
6	150	153	0.68	0.68
7	150	239	0.88	1.56
8	152	328	0.85	2.41
9	150	407	0.85	3.26
10	150	435	0.32	3.58
11	155	455	0.24	3.82
12	150	403	0.44	4.26
13	150	405	0.15	4.41
14	109	515	0.14	4.55
15	52	525	0.14	4.69

TABLE 1—*Concluded*

Kohs Block Design

C.A.	CASES	SCORE	D.V.	SCALE
5 yr.	1	0		X
6	15	2	3.10	3.10
7	27	6	1.70	4.80
8	28	11	0.90	5.70
9	34	18	1.10	6.80
10	32	25.5	0.70	7.50
11	31	30	0.50	8.00
12	37	48	0.30	8.30
13	33	61	0.30	8.60
14	20	75	0.30	8.90
15	19	89	0.50	9.40

It is obvious that for tests of approximately equal value, the matter of weighting can be safely neglected. It is taken care of by the selection of tests. Among tests of very unequal value, however, the weighting is probably of significance. An inspection of table 2 shows that the tests with the higher discriminative values, such as Kohs Block Design and the Seguin Form Board, are automatically weighted, by the present method of scale construction, so that success on either earns three times as many points as a corresponding degree of success on a test of low discriminative value, such as the Two Figure form board. In the course of time, enough tests of high discriminative value will be available, doubtless, to enable us to discard the poorer tests of the present scale, but in the meantime we need to make the best possible use of what we have at hand. By averaging the point values of the time and error scores of the Casuist form board, we avoid giving double weight to a relatively poor test.

The procedure in giving and scoring the various tests has been that described by Pintner and Paterson, by Kohs for the Block Design, and by Pintner and Anderson for the Picture Completion. It was found, however, that Sylvester's norms, based upon average and superior children, were quite useless for clinic work when his procedure was followed. Goddard's

TABLE 2

Combined Point Scale

C.A.	CUBE TEST		SEGUIN FORM BOARD		MANIKIN AND FEATURE PROFILE		TWO FIGURE FORM BOARD		CASUIST FORM BOARD				AVERAGE	MARE AND FOAL		HEALY PICTURE COMPLETION 1		KOHS BLOCK DESIGN		TOTAL
									Errors		Time									
5	2	X	37	X	3	X	200	X	DNC	X	DNC	X		107	X	89	X	0	X	X
6	4	1.60	26	2.20	4	1.0	175	0.23	30	0.84	300	0.44	0.64	71	1.18	153	0.68	2	3.10	10.63
7	5	2.60	23	3.20	4	1.0	116	0.76	12	2.02	154	1.62	1.82	62	1.67	239	1.56	6	4.80	17.41
8	6	3.93	20	4.53	5	2.33	62	1.47	8	2.49	106	2.33	2.41	48	2.65	328	2.41	11	5.70	25 43
9	6	3.93	18	5.53			47	1.91	7	2.66	93	2.58	2.62	41	3.39	407	3.26	18	6.80	29.77
10	6	3.93	16	6.53	240	3.33	47	1.91	5	3.04	78	2.91	2.98	36	4.04	435	3.58	26	7.50	33.80
11	7	5.26	15	7.09	157	4.28	38	2.28	5	3.04	68	3.26	3.15	34	4.32	455	3.82	36	8.00	38.20
12	7	5.26	14	7.76	157	4.28	38	2.28	4	3.32	66	3.34	3.33	32	4.68	493	4.26	48	8 30	40.15
13	7	5.26	12	9.09	150	4.37	29	2.91	3	3.76	66	3.34	3.55	29	5.23	505	4.41	61	8.60	43.42
14	8	6.06	11	9.76	132	4.58	29	2.91	3	3.76	58	3.71	3.74	29	5 23	515	4.55	75	8.90	45.73
15	8	6.06	11	9.76	132	4.58	29	2.91	3	3.76	58	3.71	3.74	28	5.43	525	4.69	89	9.40	46.57

procedure for this test was adopted, with more satisfactory results. The board is placed with the star on the side toward the child, for all three trials. The child sees the blocks removed. The blocks are placed in the same order each time, with the star on the bottom of the pile at the upper right corner of the board (from the child's position), the circle on the star, and the diamond on the circle. For the middle pile the cross is on the bottom, then the lozenge, then the semi-cricle, and the square on top. At the upper left corner of the board is the clipse, the triangle on that, and the oblong on top. For the Healy Picture Completion I, Dr. Florence Goodenough, then of the Minneapolis Child Guidance Clinic, and the writer agreed upon the following fixed order for the arrangement of the blocks:

(Top row): 1, bucket, 2, M. bottle, 3, wheel, 4, blank, 5, pipe, 6, shoe, 7, dog, 8, purse, 9, B. window, 10, fruit, 11, tie.

(Second row from top): 12, cup, 13, knife, 14, chicken, 15, blank, 16, hatchet, 17, blank, 18, basket, 19, blank, 20, pumpkin, 21, blank, 22, S. bird, 23, football, 24, baby.

(Third row from top): 25, baseball, 26, cat, 27, blank, 28, candle, 29, bottle, 30, hat, 31, blank, 32, fish, 33, glove, 34, mouse, 35, sprinkling can, 36, blank.

(Bottom row): 37, flowers, 38, D. cat, 39, blank, 40, F. bird, 41, books, 42, blank, 43, C. window, 44, clock, 45, scissors, 46, stool, 47, cage, 48, S. cat, 49, log, 50, cherries.

The results obtained with the above scale have more than justified expectation. Few of the tests are adequately standardized. Kohs has used only 366 cases in all. In norms for part of the tests, "five years" means 5.0, while for the others it means 5.5, and so on for the other age groups. The separate tests of the scale have been standardized upon different populations. New norms for the various tests must be obtained from a single population before this can be considered a scale in any real sense. As we have seen, norms obtained by Sylvester are being used to evaluate scores obtained with the Goddard form board and mode of procedure. One feels that the tests used must be of great inherent value to yield any worthwhile results under such conditions.

In table 3 are given the values interpolated between the year norms of table 2. These make it possible to convert total point scores into years and months of mental age.

In table 4 are given point values for all possible scores between the five year norm and the highest age norm of each test. The crude score for each test is converted into points by means of this table. These point values are totalled and converted into mental age by reading from table 2. By extrapolation, point values are assigned to scores above the highest year norm of each test. Quotients are then calculated upon a fifteen year basis. This extends the usefulness of the scale to the higher age groups. The highest score earned by a clinic patient was that of a thirteen-year-old boy. It gave him a quotient of 155. His IQ on the Kuhlmann-Binet Scale was 154.

TABLE 3

For converting point scores into years and months of mental age

YEARS	MONTHS											
	0	1	2	3	4	5	6	7	8	9	10	11
5	X	0.88	1.77	2.65	3.54	4.42	5.31	6.19	7.08	7.96	8.85	9.74
6	10.63	11.19	11.76	12.32	12.89	13.45	14.02	14.58	15.15	15.71	16.28	16.84
7	17.41	18.07	18.74	19.41	20.08	20.75	21.42	22.09	22.76	23.43	24.10	24.77
8	25.43	25.79	26.15	26.51	26.87	27.23	27.60	27.96	28.32	28.68	29.04	29.40
9	29.77	30.10	30.44	30.77	31.11	31.44	31.78	32.11	32.45	32.78	33.12	33.45
10	33.80	34.17	34.53	34.90	35.27	35.63	36.00	36.47	36.83	37.20	37.57	37.93
11	38.20	38.36	38.52	38.68	38.85	39.01	39.17	39.33	39.50	39.66	39.82	39.98
12	40.15	40.42	40.60	40.96	41.24	41.51	41.78	42.05	42.33	42.60	42.87	43.14
13	43.42	43.61	43.80	43.99	44.19	44.38	44.57	44.76	44.96	45.16	45.34	45.53
14	45.73	45.80	45.87	45.94	46.01	46.08	46.15	46.22	46.29	46.36	46.43	46.50
15	46.57											

The principles underlying the method of scale construction here employed are discussed in full in "A group point scale for the measurement of intelligence: A further study in method" (1). For the present, it is enough to state that in spite of the difference in procedure, the principles are basically the same as those of the Binet scale. For this reason it is permissible to compare results from the two scales.

TABLE 4

For determining the point value of any score

CUBE TEST (KNOX) P. & P. NORMS		
3	1	
4	1	−0.80
5 yr.	2	X
	3	0.80
6	4	1.60
7	5	2.60
8	6	3.93
9	6	3.93
10	6	3.93
11	7	4.26
12	7	4.26
13	7	4.26
14	8	5.06
15	8	5.06

+0.80 for each additional line.

SEGUIN FORM BOARD (SYLVESTER)		
5 yr.	37	X
	36	0.20
	35	0.40
	34	0.60
	33	0.80
	32	1.00
	31	1.20
	30	1.40
	29	1.60
	28	1.80
	27	2.00
6 yr.	26	2.20
	25	2.53
	24	2.87
7 yr.	23	3.20
	22	3.64
	21	4.09
8 yr.	20	4.53

SEGUIN FORM BOARD (SYLVESTER)—*Continued*		
	19	5.03
9 yr.	18	5.53
	17	6.03
10 yr.	16	6.53
11 yr.	15	7.09
12 yr.	14	7.76
	13	8.42
13 yr.	12	9.09
	11	9.76

+0.67 for each additional decrease of one second.

TWO FIGURE FORM BOARD (PINTNER)		
5 yr.	200	X
	199	0.01
	198	0.02
	197	0.03
	196	0.04
	195	0.05
	194	0.06
	193	0.07
	192	0.08
	191	0.09
	190	0.09
	189	0.10
	188	0.11
	187	0.12
	186	0.13
	185	0.14
	184	0.15
	183	0.16
	182	0.17
	181	0.18
	180	0.18

TABLE 4—*Continued*

TWO FIGURE FORM BOARD (PINTNER)—*Continued*		
	179	0.19
	178	0.20
	177	0.21
	176	0.22
6 yr.	175	0.23
	174	0.24
	173	0.25
	172	0.26
	171	0.27
	170	0.28
	169	0.29
	168	0.30
	167	0.31
	166	0.32
	165	0.32
	164	0.33
	163	0.34
	162	0.35
	161	0.30
	160	0.37
	159	0.38
	158	0.39
	157	0.40
	156	0.41
	155	0.41
	154	0.42
	153	0.43
	152	0.44
	151	0.45
	150	0.46
	140	0.47
	148	0.48
	147	0.49
	146	0.50
	145	0.50
	144	0.51
	143	0.52
	142	0.53
	141	0.54
	140	0.55
	139	0.56
	138	0.57
	137	0.58
	136	0.59
	135	0.59
	134	0.60
	133	0.61
	132	0.62
	131	0.63
	130	0.64
	120	0.65
	128	0.66
	127	0.67
	126	0.68
	125	0.68
	124	0.69
	123	0.70
	122	0.71
	121	0.72
	120	0.73
	110	0.74
	118	0.75
	117	0.76
7 yr.	116	0.70
	115	0.77
	114	0.78
	113	0.80
	112	0.81
	111	0.82
	110	0.84
	109	0.85
	108	0.86
	107	0.88
	106	0.89
	105	0.90
	104	0.92
	103	0.93
	102	0.94
	101	0.96
	100	0.97

TABLE 4—*Continued*

TWO FIGURE FORM BOARD (PINTNER)—*Continued*		
	99	0.98
	98	1.00
	97	1.01
	96	1.02
	95	1.04
	94	1.05
	93	1.06
	92	1.08
	91	1.09
	90	1.10
	89	1.12
	88	1.13
	87	1.14
	86	1.16
	85	1.17
	84	1.18
	83	1.20
	82	1.21
	81	1.22
	80	1.24
	79	1.25
	78	1.26
	77	1.28
	76	1.29
	75	1.30
	74	1.32
	73	1.33
	72	1.34
	71	1.36
	70	1.37
	69	1.38
	68	1.40
	67	1.41
	66	1.42
	65	1.44
	64	1.45
	63	1.46
8 yr.	62	1.47
	61	1.50
	60	1.53

TWO FIGURE FORM BOARD (PINTNER)—*Continued*		
	59	1.56
	58	1.59
	57	1.62
	56	1.65
	55	1.68
	54	1.71
	53	1.74
	52	1.77
	51	1.80
	50	1.83
	49	1.86
	48	1.89
9 yr.	47	1.91
	46	1.95
	45	1.99
	44	2.03
	43	2.07
	42	2.11
	41	2.15
	40	2.19
	39	2.23
11 yr.	38	2.28
	37	2.35
	36	2.41
	35	2.48
	34	2.55
	33	2.62
	32	2.69
	31	2.76
	30	2.83
13 yr.	29	2.91
	28	2.98
	27	3.05
	26	3.12
	25	3.19

+0.07 for each additional decrease of one second in time.

TABLE 4—*Continued*

CASUIST FORM BOARD (KNOX) TIME		
5 yr.	DNC	X
6 yr.	300	0.44
	299	0.45
	298	0.46
	297	0.47
	296	0.48
	295	0.48
	294	0.49
	293	0.50
	292	0.51
	291	0.52
	290	0.52
	289	0.53
	288	0.54
	287	0.55
	286	0.56
	285	0.56
	284	0.57
	283	0.58
	282	0.59
	281	0.60
	280	0.60
	279	0.61
	278	0.62
	277	0.63
	276	0.64
	275	0.64
	274	0.65
	273	0.66
	272	0.67
	271	0.68
	270	0.68
	269	0.69
	268	0.70
	267	0.71
	266	0.72
	265	0.72
	264	0.73
	263	0.74
	262	0.75
	261	0.76
	260	0.76
	259	0.77
	258	0.78
	257	0.79
	256	0.80
	255	0.80
	254	0.81
	253	0.82
	252	0.83
	251	0.84
	250	0.84
	249	0.85
	248	0.86
	247	0.87
	246	0.88
	245	0.88
	244	0.89
	243	0.90
	242	0.91
	241	0.92
	240	0.92
	239	0.93
	238	0.94
	237	0.95
	236	0.96
	235	0.96
	234	0.97
	233	0.98
	232	0.99
	231	1.00
	230	1.00
	229	1.01
	228	1.02
	227	1.03
	226	1.04
	225	1.04
	224	1.05
	223	1.06
	222	1.07

TABLE 4—*Continued*

	CASUIST FORM BOARD (KNOX) TIME—*Continued*	
	221	1.08
	220	1.08
	219	1.09
	218	1.10
	217	1.11
	216	1.12
	215	1.12
	214	1.13
	213	1.14
	212	1.15
	211	1.16
	210	1.16
	209	1.17
	208	1.18
	207	1.19
	206	1.20
	205	1.20
	204	1.21
	203	1.22
	202	1.23
	201	1.24
	200	1.24
	199	1.25
	198	1.26
	197	1.27
	196	1.28
	195	1.28
	194	1.29
	193	1.30
	192	1.31
	191	1.32
	190	1.32
	189	1.33
	188	1.34
	187	1.35
	186	1.36
	185	1.36
	184	1.37
	183	1.38
	182	1.39
	181	1.40
	180	1.40
	179	1.41
	178	1.42
	177	1.43
	176	1.44
	175	1.44
	174	1.45
	173	1.46
	172	1.47
	171	1.48
	170	1.48
	169	1.49
	168	1.50
	167	1.51
	166	1.52
	165	1.52
	164	1.53
	163	1.54
	162	1.55
	161	1.56
	160	1.56
	159	1.57
	158	1.58
	157	1.59
	156	1.60
	155	1.61
7 yr.	154	1.62
	153	1.63
	152	1.64
	151	1.66
	150	1.68
	149	1.69
	148	1.70
	147	1.71
	146	1.72
	145	1.74
	144	1.76

TABLE 4—*Continued*

	CASUIST FORM BOARD (KNOX) TIME—*Continued*	
	143	1.77
	142	1.78
	141	1.79
	140	1.81
	139	1.83
	138	1.84
	137	1.85
	136	1.86
	135	1.88
	134	1.90
	133	1.91
	132	1.92
	131	1.93
	130	1.95
	129	1.97
	128	1.98
	127	1.99
	126	2.00
	125	2.02
	124	2.04
	123	2.05
	122	2.06
	121	2.07
	120	2.09
	119	2.11
	118	2.12
	117	2.13
	116	2.14
	115	2.16
	114	2.18
	113	2.19
	112	2.20
	111	2.21
	110	2.23
	109	2.26
	108	2.29
	107	2.31
8 yr.	106	2.33
	105	2.35
	104	2.37
	103	2.39
	102	2.41
	101	2.43
	100	2.45
	99	2.47
	98	2.49
	97	2.51
	96	2.52
	95	2.54
	94	2.56
9 yr.	93	2.58
	92	2.60
	91	2.62
	90	2.64
	89	2.66
	88	2.69
	87	2.71
	86	2.73
	85	2.75
	84	2.77
	83	2.80
	82	2.82
	81	2.84
	80	2.86
	79	2.89
10 yr.	78	2.91
	77	2.94
	76	2.98
	75	3.01
	74	3.05
	73	3.08
	72	3.12
	71	3.15
	70	3.19
	69	3.22
11 yr.	68	3.26
	67	3.30
12 yr.	66	3.34

TABLE 4—*Continued*

CASUIST FORM BOARD (KNOX) TIME—*Continued*

	65	3.38
	64	3.43
	63	3.47
	62	3.52
	61	3.56
	60	3.61
	59	3.66
14 yr.	58	3.71

+0.046 for each additional decrease of one second.

CASUIST FORM BOARD (ERRORS)

5 yr.	DNC	X
6 yr.	30	0.84
	29	0.90
	28	0.97
	27	1.03
	26	1.10
	25	1.16
	24	1.23
	23	1.29
	22	1.36
	21	1.42
	20	1.49
	19	1.55
	18	1.62
	17	1.68
	16	1.75
	15	1.81
	14	1.88
	13	1.95
7 yr.	12	2.02
	11	2.14
	10	2.26
	9	2.38
8 yr.	8	2.49

CASUIST FORM BOARD (ERRORS)—*Continued*

9 yr.	7	2.66
	6	2.85
10 yr.	5	3.04
12 yr.	4	3.32
13 yr.	3	3.76
	2	4.20
	1	4.64

MANIKIN

3 yr.	0	−1.50
	1	−1.00
4 yr.	2	−0.50
5 yr.	3	X
6 yr.	4	1.–
8 yr.	5	2.33

FEATURE PROFILE

9 yr.	DNC	−0.01
10 yr.	240	3.33
	239	3.34
	238	3.35
	237	3.36
	236	3.37
	235	3.39
	234	3.40
	233	3.41
	232	3.42
	231	3.44
	230	3.45
	229	3.46
	228	3.47
	227	3.48
	226	3.49
	225	3.51

TABLE 4—*Continued*

	FEATURE PROFILE—*Continued*	
	224	3.52
	223	3.53
	222	3.54
	221	3.56
	220	3.57
	219	3.58
	218	3.59
	217	3.60
	216	3.61
	215	3.63
	214	3.64
	213	3.65
	212	3.66
	211	3.68
	210	3.69
	209	3.70
	208	3.71
	207	3.72
	206	3.73
	205	3.74
	204	3.75
	203	3.76
	202	3.77
	201	3.79
	200	3.80
	199	3.81
	198	3.82
	197	3.83
	196	3.84
	195	3.85
	194	3.86
	193	3.87
	192	3.88
	191	3.90
	190	3.91
	189	3.92
	188	3.93
	187	3.94
	186	3.95
	185	3.96
	184	3.97
	183	3.98
	182	3.99
	181	4.01
	180	4.02
	179	4.03
	178	4.04
	177	4.05
	176	4.06
	175	4.07
	174	4.08
	173	4.09
	172	4.10
	171	4.11
	170	4.12
	169	4.14
	168	4.15
	167	4.16
	166	4.17
	165	4.18
	164	4.19
	163	4.20
	162	4.21
	161	4.22
	160	4.24
	159	4.25
	158	4.26
11 yr.	157	4.28
	156	4.29
	155	4.30
	154	4.32
	153	4.33
	152	4.34
	151	4.36
13 yr.	150	4.37
	149	4.38
	148	4.39
	147	4.40
	146	4.41
	145	4.42

TABLE 4—*Continued*

FEATURE PROFILE—*Continued*		
	144	4.44
	143	4.45
	142	4.46
	141	4.47
	140	4.48
	139	4.49
	138	4.51
	137	4.52
	136	4.53
	135	4.54
	134	4.55
	133	4.56
14 yr.	132	4.58
	131	
	130	

$+0.01\frac{1}{6}$ for each additional decrease of one second.

MARE AND FOAL (HEALY)		
5 yr.	107	X
	106	0.03
	105	0.06
	104	0.10
	103	0.13
	102	0.16
	101	0.20
	100	0.23
	99	0.26
	98	0.29
	97	0.32
	96	0.35
	95	0.39
	94	0.42
	93	0.45
	92	0.49
	91	0.52
	90	0.55
	89	0.59

MARE AND FOAL (HEALY)—*Continued*		
	88	0.62
	87	0.65
	86	0.69
	85	0.72
	84	0.75
	83	0.79
	82	0.82
	81	0.85
	80	0.89
	79	0.92
	78	0.95
	77	0.99
	76	1.02
	75	1.05
	74	1.09
	73	1.12
	72	1.15
6 yr.	71	1.18
	70	1.23
	69	1.29
	68	1.34
	67	1.40
	66	1.45
	65	1.51
	64	1.56
	63	1.61
7 yr.	62	1.67
	61	1.74
	60	1.81
	59	1.87
	58	1.94
	57	2.01
	56	2.08
	55	2.15
	54	2.22
	53	2.29
	52	2.36
	51	2.43
	50	2.50
	49	2.57

TABLE 4—*Continued*

MARE AND FOAL (HEALY)—*Continued*		
8 yr.	48	2.65
	47	2.76
	46	2.86
	45	2.97
	44	3.07
	43	3.18
	42	3.28
9 yr.	41	3.39
	40	3.52
	39	3.65
	38	3.78
	37	3.91
10 yr.	36	4.04
	35	4.18
11 yr.	34	4.32
	33	4.50
	32	4.68
	31	4.86
	30	5.04
13 yr.	29	5.23
15 yr.	28	5.43
	27	5.63
	26	5.83
	25	6.03
	24	6.23
	23	6.43
	22	6.63
	21	6.83
	20	7.03
	19	7.23
	18	7.43
	17	7.63

+0.20 for each additional decrease of one second.

HEALY PICTURE COMPLETION I		
5 yr.	89	X
	90	0.01
	91	0.02
	92	0.03
	93	0.04
	94	0.05
	95	0.06
	96	0.07
	97	0.08
	98	0.09
	99	0.11
	100	0.12
	101	0.13
	102	0.14
	103	0.15
	104	0.16
	105	0.17
	106	0.18
	107	0.19
	108	0.21
	109	0.22
	110	0.23
	111	0.24
	112	0.25
	113	0.26
	114	0.27
	115	0.28
	116	0.29
	117	0.30
	118	0.31
	119	0.32
	120	0.33
	121	0.34
	122	0.35
	123	0.36
	124	0.37
	125	0.38
	126	0.39
	127	0.40
	128	0.41

TABLE 4—*Continued*

HEALY PICTURE COMPLETION I—*Continued*		
	129	0.42
	130	0.43
	131	0.44
	132	0.45
	133	0.46
	134	0.47
	135	0.48
	136	0.49
	137	0.51
	138	0.52
	139	0.53
	140	0.54
	141	0.55
	142	0.56
	143	0.57
	144	0.58
	145	0.59
	146	0.61
	147	0.62
	148	0.63
	149	0.64
	150	0.65
	151	0.66
	152	0.67
6 yr.	153	0.68
	154	0.69
	155	0.70
	156	0.71
	157	0.72
	158	0.73
	159	0.74
	160	0.75
	161	0.76
	162	0.77
	163	0.78
	164	0.79
	165	0.80
	166	0.81
	167	0.82
	168	0.83
	169	0.84
	170	0.85
	171	0.86
	172	0.87
	173	0.88
	174	0.89
	175	0.90
	176	0.91
	177	0.92
	178	0.93
	179	0.94
	180	0.95
	181	0.96
	182	0.97
	183	0.98
	184	0.99
	185	1.00
	186	1.01
	187	1.02
	188	1.03
	189	1.04
	190	1.05
	191	1.06
	192	1.07
	193	1.08
	194	1.09
	195	1.10
	196	1.11
	197	1.12
	198	1.13
	199	1.14
	200	1.16
	201	1.17
	202	1.18
	203	1.19
	204	1.20
	205	1.21
	206	1.22
	207	1.23
	208	1.24
	209	1.25

TABLE 4—*Continued*

HEALY PICTURE COMPLETION I—*Continued*		
	210	1.26
	211	1.27
	212	1.28
	213	1.29
	214	1.30
	215	1.31
	216	1.32
	217	1.33
	218	1.34
	219	1.35
	220	1.36
	221	1.37
	222	1.38
	223	1.39
	224	1.40
	225	1.41
	226	1.42
	227	1.43
	228	1.44
	229	1.45
	230	1.46
	231	1.47
	232	1.48
	233	1.49
	234	1.50
	235	1.51
	236	1.52
	237	1.53
	238	1.54
7 yr.	239	1.56
	240	1.57
	241	1.58
	242	1.59
	243	1.60
	244	1.61
	245	1.62
	246	1.63
	247	1.64
	248	1.65
	249	1.66
	250	1.67
	251	1.68
	252	1.69
	253	1.70
	254	1.71
	255	1.72
	256	1.73
	257	1.74
	258	1.75
	259	1.76
	260	1.77
	261	1.77
	262	1.78
	263	1.79
	264	1.80
	265	1.81
	266	1.82
	267	1.83
	268	1.84
	269	1.85
	270	1.86
	271	1.87
	272	1.88
	273	1.89
	274	1.90
	275	1.91
	276	1.92
	277	1.93
	278	1.94
	279	1.95
	280	1.96
	281	1.96
	282	1.97
	283	1.98
	284	1.99
	285	2.00
	286	2.01
	287	2.02
	288	2.03
	289	2.04
	290	2.05

TABLE 4—*Continued*

	HEALY PICTURE COMPLETION I—*Continued*	
	291	2.06
	292	2.07
	293	2.08
	294	2.09
	295	2.10
	296	2.11
	297	2.12
	298	2.13
	299	2.14
	300	2.15
	301	2.15
	302	2.16
	303	2.17
	304	2.18
	305	2.19
	306	2.20
	307	2.21
	308	2.22
	309	2.23
	310	2.24
	311	2.25
	312	2.26
	313	2.27
	314	2.28
	315	2.29
	316	2.30
	317	2.31
	318	2.32
	319	2.33
	320	2.34
	321	2.35
	322	2.36
	323	2.37
	324	2.38
	325	2.39
	326	2.40
	327	2.41
8 yr.	328	2.41
	329	2.42
	330	2.43
	331	2.44
	332	2.45
	333	2.46
	334	2.47
	335	2.48
	336	2.49
	337	2.50
	338	2.51
	339	2.52
	340	2.53
	341	2.55
	342	2.56
	343	2.57
	344	2.58
	345	2.59
	346	2.60
	347	2.61
	348	2.62
	349	2.63
	350	2.64
	351	2.65
	352	2.66
	353	2.67
	354	2.69
	355	2.70
	356	2.71
	357	2.72
	358	2.73
	359	2.74
	360	2.75
	361	2.76
	362	2.77
	363	2.78
	364	2.79
	365	2.80
	366	2.81
	367	2.83
	368	2.84
	369	2.85
	370	2.86
	371	2.87

TABLE 4—*Continued*

HEALY PICTURE COMPLETION I—*Continued*			HEALY PICTURE COMPLETION I—*Continued*		
	372	2.88		412	3.31
	373	2.89		413	3.32
	374	2.90		414	3.34
	375	2.91		415	3.35
	376	2.92		416	3.36
	377	2.93		417	3.37
	378	2.94		418	3.38
	379	2.95		419	3.39
	380	2.97		420	3.40
	381	2.98		421	3.42
	382	2.99		422	3.43
	383	3.00		423	3.44
	384	3.01		424	3.45
	385	3.02		425	3.46
	386	3.03		426	3.47
	387	3.04		427	3.48
	388	3.05		428	3.50
	389	3.06		429	3.51
	390	3.07		430	3.52
	391	3.08		431	3.53
	392	3.09		432	3.54
	393	3.11		433	3.55
	394	3.12		434	3.56
	395	3.13	10 yr.	435	3.58
	396	3.14		436	3.59
	397	3.15		437	3.60
	398	3.16		438	3.61
	399	3.17		439	3.62
	400	3.18		440	3.64
	401	3.19		441	3.65
	402	3.20		442	3.66
	403	3.21		443	3.67
	404	3.22		444	3.68
	405	3.23		445	3.70
	406	3.25		446	3.71
9 yr.	407	3.26		447	3.72
	408	3.27		448	3.73
	409	3.28		449	3.74
	410	3.29		450	3.76
	411	3.30		451	3.77

TABLE 4—*Continued*

HEALY PICTURE COMPLETION I—*Continued*		
	452	3.78
	453	3.79
	454	3.80
11 yr.	455	3.82
	456	3.83
	457	3.84
	458	3.85
	459	3.86
	460	3.87
	461	3.89
	462	3.90
	463	3.91
	464	3.02
	465	3.93
	466	3.94
	467	3.96
	468	3.97
	469	3.98
	470	3.99
	471	4.00
	472	4.01
	473	4.03
	474	4.04
	475	4.05
	476	4.06
	477	4.07
	478	4.08
	479	4.10
	480	4.11
	481	4.12
	482	4.13
	483	4.14
	484	4.15
	485	4.17
	486	4.18
	487	4.19
	488	4.20
	489	4.21
	490	4.22
	491	4.23

HEALY PICTURE COMPLETION I—*Continued*		
	491	4.23
	492	4.24
12 yr.	493	4.26
	494	4.27
	495	4.28
	496	4.29
	497	4.31
	498	4.32
	499	4.33
	500	4.34
	501	4.36
	502	4.37
	503	4.38
	504	4.39
13 yr.	505	4.41
	506	4.42
	507	4.43
	508	4.45
	509	4.46
	510	4.48
	511	4.49
	512	4.50
	513	4.52
	514	4.53
14 yr.	515	4.55
	516	4.56
	517	4.57
	518	4.59
	519	4.60
	520	4.62
	521	4.63
	522	4.64
	523	4.66
	524	4.67
15 yr.	526	4.60

+0.014 for each additional crude score unit.

TABLE 4—*Continued*

KOHS BLOCK DESIGN		
5 yr.	0	X
	1	1.55
6 yr.	2	3.10
	3	3.53
	4	3.95
	5	4.38
7 yr.	6	4.80
	7	4.98
	8	5.16
	9	5.34
	10	5.52
8 yr.	11	5.70
	12	5.86
	13	6.01
	14	6.17
	15	6.33
	16	6.49
	17	6.64
9 yr.	18	6.80
	19	6.90
	20	7.00
	21	7.10
	22	7.20
	23	7.30
	24	7.40
10 yr.	25	7.50
	26	7.54
	27	7.59
	28	7.63
	29	7.68
	30	7.72
	31	7.77
	32	7.81
	33	7.86
	34	7.90
	35	7.95
KOHS BLOCK DESIGN—*Continued*		
11 yr.	36	8.00
	37	8.02
	38	8.05
	39	8.07
	40	8.10
	41	8.12
	42	8.15
	43	8.17
	44	8.20
	45	8.22
	46	8.25
	47	8.27
12 yr.	48	8.30
	49	8.32
	50	8.34
	51	8.37
	52	8.39
	53	8.41
	54	8.44
	55	8.46
	56	8.48
	57	8.51
	58	8.53
	59	8.55
	60	8.58
13 yr.	61	8.60
	62	8.62
	63	8.64
	64	8.66
	65	8.68
	66	8.70
	67	8.72
	68	8.74
	69	8.76
	70	8.78
	71	8.81
	72	8.83
	73	8.85

TABLE 4—*Continued*

KOHS BLOCK DESIGN—*Continued*		
	74	8.87
14 yr.	75	8.90
	76	8.94
	77	8.98
	78	9.02
	79	9.06
	80	9.10
	81	9.14
	82	9.18
	83	9.22
	84	9.25
	85	9.28
	86	9.31
	87	9.34
	88	9.37
15 yr.	89	9.40
	90	9.47
	91	9.55
	92	9.63
	93	9.70
	94	9.78
	95	9.86
	96	9.93
	97	10.01
	98	10.09
	99	10.16
	100	10.24
	101	10.32
16 yr.	102	10.40
	103	10.56
	104	10.72
	105	10.88
	106	11.04
	107	11.20
	108	11.36
	109	11.52
	110	11.68
	111	11.84
17 yr.	112	12.00
	113	12.13
	114	12.26
	115	12.40
	116	12.53
	117	12.66
	118	12.80
	119	12.93
	120	13.06
18 yr.	121	13.20
	122	13.50
	123	13.80
	124	14.1
	125	14.4
	126	14.7
19 yr.	127	15.00

+0.3 for each additional point.

For 46 clinic subjects, IQ's are available for both the Kuhlmann-Binet and new point performance scales. To find a measure that would be stable whether the regression was linear or not, and whatever the range of values considered, the probable error of measurement was calculated by Nygaard's (5) formula based on averages. It was found to be 4.39. This is much larger than that obtained between 275 Kuhlmann-Binet IQ's and those of the group point scale above referred to, which was 2.60, but is surprisingly good considering that the P.E.

of the Stanford-Binet scale, as published by Terman (6), is 4.50.

It would seem, then, that a performance scale, when constructed according to Binet principles, can be trusted to do the work of a Binet scale in measuring general intelligence. In its present tentative state of standardization, one would never use it in place of a Binet measure except form necessity. For the deaf child, however, and for the child with speech difficulty, whether the difficulty is one of emotional repression, foreign language handicap or so-called speech defect, it is highly useful in furnishing an intelligence rating that is far more independent of training than are any of the written tests.

The flexibility of the scale adds much to its practical value for clinic use. Separate tests are not infrequently omitted for one reason or another, in giving the scale. In such cases the point value of the omitted test is subtracted from the combined scale norms between which the obtained point score would fall, and the distance above the lower norm is converted into a fraction of a year.

Plans are under way for adequate standardization of the present scale as a whole. The Porteus Maze test will be included, and should increase the value of the scale. It is possible that both form boards will be omitted. A better test in place of the Manikin—Feature Profile would be highly desirable. It is believed that the reliability of the Knox Cube test will be increased by giving it both at the beginning and at the end of the test series.

REFERENCES

(1) Arthur and Woodrow: An Absolute Intelligence Scale. Journal of Applied Psychology, Vol. III, June, 1919.

(2) Pintner and Paterson: A Scale of Performance Tests. D. Appleton & Company, New York, 1923.

(3) Kohs, S. C.: Intelligence Measurement. Macmillan, New York, 1923.

(4) Pintner and Anderson: The Picture Completion Test. Warwick & York, Baltimore, 1917.

(5) Nygaard, P. H.: Advantages of the P.E. of measurement. Journal of Educational Psychology, 1923.

(6) Terman, Lewis M. Intelligence of School Children, Houghton Mifflin, Boston, 1919, p. 142.

A NOTE ON MOTOR ACTIVITY AS CONDITIONED BY EMOTIONAL STATES[1]

H. H. REMMERS AND L. A. THOMPSON, JR.

Purdue University

There is scriptural as well as popular warrant for the assertion that there is a positive relationship between pleasurable emotional states and motor activity. To "jump for joy" and to "dance and make a joyful noise before the Lord" obviously are, from a psychological viewpoint, similar if not identical phenomena. Aside from these two weighty authorities it is of practical as well as theoretical importance to measure quantitatively this relationship to ascertain whether science agrees or disagrees with the dicta of Demos and the Bible. To answer the question with reference to a relatively controlled situation was the purpose of the experiment here described.

PROCEDURE

The data were gathered from sophomore engineering students in a beginning course in psychology. The method followed in most essentials the procedure outlined by Langfeld and Allport.[2] The treatment of the quantitative results and the empirical checks described below are the only features of this study for which any originality is claimed. The directions given to students will make clear the procedure.

---

[1] Responsibility for this study is divided as follows: Thompson and Remmers together collected the data of the main experiment. These data were summarized by Remmers. Thompson suggested the empirical check (see context) and gathered and summarized the data for this. For the statistical treatment and interpretation of all the data and the writing of this article Remmers is wholly responsible.

[2] Langfeld, H. S., and Allport, F. H., An Elementary Course in Laboratory Psychology, p. 146 ff.

DIRECTIONS

You are to think of a very pleasant experience that you have had. When you have reinstated the emotional state as nearly as possible, you will draw free-hand a line of from 2 to 5 inches in length approximately at the middle of the unruled white sheet I have given you. This is to avoid the possibility of your having an objective criterion of judging successive lengths of lines to be drawn later. When you have drawn the line near the top of the middle of the sheet, fold the sheet over in such a manner that you do not see the first line you drew. Now think of a very unpleasant experience. While thinking of this try to draw another line equal in length to the one covered up. Now think of another unpleasant experience. Cover the second line in the same way as the first, attempt again to draw a line of the same length. I have placed on the board a scheme of "double fatigue" procedure in which P = pleasant and U = unpleasant. Note that it goes like this:

P
U
U
P
P
U
etc.

You are to continue this for a total of forty lines. Call the successive pairs A_1, pleasant; A_2, unpleasant; B_2, unpleasant; B_1, pleasant; etc. Use four or five different emotional experiences of each type if you can think of that many.

Before beginning the experiment the instructions were again carefully gone over and an opportunity given for asking questions on any part of the procedure not perfectly clear. When it was quite certain that all understood, the experiment proceeded.

The successive pairs of lines were then measured in millimeters, tabulated, and summarized.

Limitations of space forbid the inclusion of the totals of individual subjects. Suffice it to say they varied for the various categories at the head of the columns as noted in table 1.

As is obvious even from this table, individual variations are large and may go in either direction under the conditions of the experiment. In other words, for some subjects pleasant vs. unpleasant moods produce a greater total length of lines while for others the results are reversed. The query immediately presents itself: are these fluctuations not a matter of "chance?"[3] The answer in terms of all the data gathered is negative. Table 2 gives the essential facts.

The observed difference of 4329 mm. between the two totals is greater than chance would account for. We may reasonably assume that the two distributions from which these totals are computed are amenable to the criterion of the Bernoulli theorem. The probable error of such a dichotomous division as our data involve is given by the formula.[4]

TABLE 1

Totals of Individual subjects

	LENGTH OF PLEASANT LINES	LENGTH OF UNPLEASANT LINES	AMOUNT PLEASANT LINES LONGER	PER CENT PLEASANT LINES LONGER	AMOUNT UNPLEASANT LINES LONGER	PER CENT UNPLEASANT LINES LONGER
Maximum.........	4,285	4,220	300	453.4	722	162.4
Minimum.........	511	536	6	4.7	0	0

$$\text{P.E.} = 0.6745 \sqrt{\frac{p\,q}{n}}$$

where P.E. = probable error

p = probability of success (i.e., pleasant longer)

q = probability of failure (i.e., unpleasant longer)

$p + q = 1$

n = number of cases

Substituting and solving this equation for the values of pleasant and unpleasant length of lines we obtain

$$\text{P.E.} = 0.6745 \sqrt{\frac{(0.509)\ (0.491)}{276167}}$$

$$= 0.00064$$

---

[3] *Chance* is defined for purposes of this discussion as *ignorance of cause*—the meaning it must be given in a universe viewed from the standpoint of science.

[4] See Rietz, H. L. et al., A Handbook of Mathematical Statistics, p. 74.

This may be interpreted by stating that it is an even wager that the true value of p in terms of percentage is 0.509 ± 0.00064 and that this P.E. indicates that the observed difference could never occur by chance. The odds in favor of this difference being significant runs into billions.[5]

TABLE 2

Totals and averages for all subjects. N = 84

	LENGTH OF PLEASANT LINES	LENGTH OF UNPLEASANT LINES	AMOUNT PLEASANT LINES LONGER	PER CENT PLEASANT LINES LONGER	AMOUNT UNPLEASANT LINES LONGER	PER CENT UNPLEASANT LINES LONGER
	mm.	mm.				
	140,248	135,919	7,895	9,134.2	3,566	4,129.9
Average per subject..	1,669.62	1,618.08	94.0	108.70	42.40	49.17
Average per line......	83.481	80.904	4.70	5.435	2.123	2.459

TABLE 3

Totals and averages for 51 subjects. Chance lines

LENGTH OF LINE A	LENGTH OF LINE B	LONGER A	PER CENT LONGER A	LONGER B	PER CENT LONGER B
81,866	81,038	2,209	2,733.04	1,755	2,130.67
80.25	79.45	2.15	2.67	1.72	2.25

AN EMPIRICAL CHECK EXPERIMENT

A further sampling of lengths of lines under identical conditions except for the variables of pleasant and unpleasant moods was suggested by Mr. Thompson and was undertaken by him. He used as subjects 51 of the original 84 subjects. They were told to draw forty lines as before, but with no other conditions. Both halves of these were measured in the same order as in the other experiment and were labelled Lines A_1 and B_2 respectively. The results are given in table 3. Applying our criterion of P.E. we obtain

$$\text{P.E.} = 0.6745\sqrt{\frac{(0.50025)\ (0.49975)}{162904}} = 0.00083$$

---

[5] *Ibid.*, p. 100.

That is, the difference between p and q is less than 1 P.E. and therefore a true chance difference. The assumptions made, therefore, are corroborated by the empirical check.

A SECOND CHECK

Still another check was undertaken[6] by making a frequency distribution of a sample of forty-three individual results. The data are shown in table 4.

TABLE 4

Frequency distribution of pleasant and unpleasant lines. $N = 43$

LENGTH OF LINES	FREQUENCIES	
	Pleasant	Unpleasant
mm.		
25–34		11
35–44	18	22
45–54	84	103
55–64	130	123
65–74	104	108
75–84	123	93
85–94	76	101
95–104	106	93
105–114	62	70
115–124	59	57
125–134	29	45
135–144	20	18
145–154	19	12
155–164	14	4
165–174	1	0
Totals	860	860
Mean	86.78	83.47
σ	28.5	27.7
σ_M	0.972	0.945
Diff.$_M$	3.31 ± 0.914	

It is apparent from the P.E. of the difference of the means that although not as favorable to our conclusions as the application of the Bernoulli theorem, the conclusion as to the significance of the difference is still practical certainty.[7]

---

[6] For the statistical labor of this I am indebted to my wife.—H. H. R.

SOME PRACTICAL AND THEORETICAL CONSIDERATIONS

One of the questions that comes to mind is whether sound pedagogy sanctions the inclusion of this experiment in an elementary manual of laboratory exercises in psychology.[8] The students are asked to interpret results which are likely to disagree as between any two individuals, not only as to relative differences of length of lines but also as to which kind of lines, pleasant or unpleasant, shall be longer. They are asked to do this interpretation, moreover, when they have not become acquainted with statistical techniques adequate to the job. The net result is likely to be a contempt on the part of the undiscriminating student for the subject matter of psychology as a whole, on the basis of the well known mechanism of "associative shifting," to use Thorndike's terminology. They are likely to "throw out the child with the bath."

Another question may be raised as to the causes for the observed individual differences. May the data be considered as supporting any of the various "type" theories of emotion, whether of Doris Blake, Mr. McCutcheon's "Glads" and "Blues" or the ancient classification of temperaments into sanguine-choleric and melancholic-phlemgatic? The answer is no. The variations in results, of which a tabulation might be made, show that there are all degrees of variation from pure chance differences, i.e., no differences, to the variations shown in table 1.

A partial explanation at least is to be found, I believe, in the vague meanings attached to "pleasant" and "unpleasant." Several of the subjects volunteered the information that the unpleasant experiences in their case were anger-evoking experiences and these produced longer lines. Introspective analysis and more carefully defined categories of emotion would probably be illuminating in this connection.

---

[7] For a definition of "practical certainty," see Herring, Jour. Educ. Psych., December, 1924.

[8] Criticism of this particular experiment is by no means to be extended to the rest of Langfeld and Allport's Manual. I have used much of their material with excellent results. In my judgment the value of the Manual would be increased by the inclusion of at least tentative norms of experimental results.

SOCIAL STATUS OF THE CLERICAL WORKER AND HIS PERMANENCE ON THE JOB

MARION A. BILLS

Life Insurance Sales Research Bureau, Hartford, Connecticut

During the year 1922, 59 clerical workers were hired in a life insurance company of medium size. Each of the 59 filled out an application blank and on it stated among other things, the occupation of the father.

These parental occupations have been studied and grouped into seven classes:

I. Unskilled laborer
II. Semi-skilled mechanical worker
III. Skilled mechanical worker
IV. Clerical worker
V. Salesman or manager of a department in a company
VI. Proprietor of a small business
VII. Professional man (this group would include managers or owners of large business concerns but there were none in the 59 cases studied)

The classes are given above in the order which, we believe, the average clerk would consider that of their social status. The order of grouping is not that of financial returns. The salary of the skilled mechanical worker and even of the semi-skilled would in many cases exceed that of the clerical worker; however, anyone who has known a large number of clerical workers, especially those employed in an office that is connected with a factory, will, we believe, agree that women, particularly, consider a clerical job to carry a higher social standing than a mechanical job.

Two occupations only, out of the whole number of parental occupations reported, offered difficulty of classification. The first was a policeman, the second, a customs inspector. On account of the salary received by the policeman, we placed him with the skilled mechanic group. The customs inspector was dropped from the list since we were unable to determine his specific duties or salary.

Figure 1 shows the number of clerks whose fathers fall into each of these occupational groups. It will be noted that the largest number are found in the group of skilled workmen and that, considering the small number of cases involved, a fairly regular frequency distribution curve is formed, around that class as a mode.

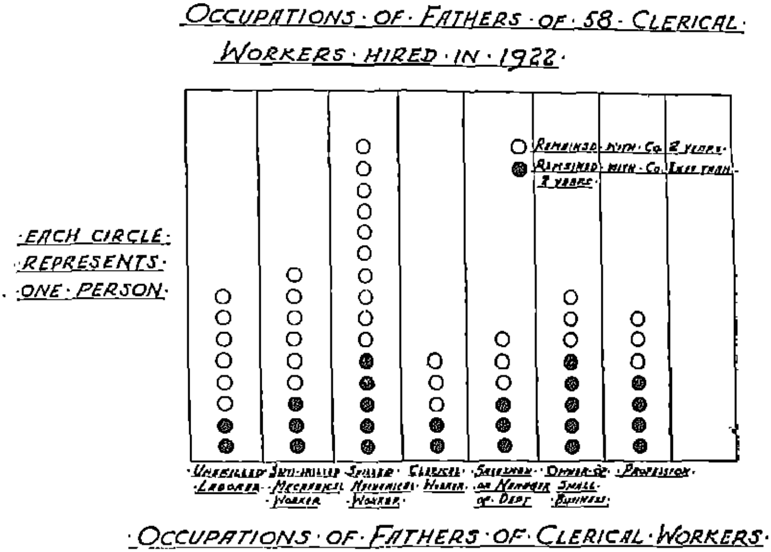


FIG. 1

At the end of 1924 an examination was made to see how many of those entering in 1922 had remained with the company through their second anniversary. It was found that 24 of the 58 employed had left the company, 34 still remaining. That is, there had been in the two years a 41 per cent turnover, or an average turnover of 20.5 per cent each year, a record which all personnel managers would consider very good.

The selection of these 58 workers had been very carefully made on the basis of intelligence and general ability, tests having been given for both general intelligence and trade proficiency. Practically the entire turnover had been voluntary on the part of the employees.

The distribution of this turnover, by the occupations of the fathers, shows an interesting result. This is also shown in figure 1. We have indicated those remaining with the company by an open circle and those leaving the company by a filled in circle. Table 1 gives the same information.

TABLE 1

OCCUPATION OF FATHER	NUMBER OF CLERKS REMAINING FOR TWO YEARS OR MORE	NUMBER OF CLERKS REMAINING LESS THAN TWO YEARS	PER CENT OF CLASS REMAINING LESS THAN TWO YEARS
			per cent
I. Unskilled worker	6	2	25
II. Semi-skilled mechanical worker	6	3	33
III. Skilled mechanical worker	10	5	33
IV. Clerical worker	3	2	40
V. Salesman and department manager	3	3	50
VI. Proprietor of small business	3	5	02
VII. Profession	3	4	57
Total	34	24	41

It is noteworthy that considering the small number of cases involved there is a remarkably consistent increase in the per cent leaving, from class I whose fathers were unskilled laborers to classes VI and VII whose fathers were professional men or proprietors of a small business. The turnover here ranges from 25 to 62 per cent.

A similar study of the employment records of two other companies, under less strict conditions, however, shows the same facts concerning the variation in the length of service of clerical workers from different social strata.

The statistical data are, of course, exceedingly meager, but nevertheless are suggestive. An explanation of the tendency readily suggests itself. If the clerical worker is bettering himself socially by taking a clerical job and continuing in it, he tends to remain with the company, but if the clerical job means a demotion in his social status, he tends to leave the job.

Another explanation also suggests itself. A boy entering the clerical field from the family of an unskilled laborer feels, as he looks forward to possible promotion some day to section or division head at $3000 a year, that the future is very promising. On the other hand, a boy from the family of a salesman or a manager of a small business, does not look upon such a prospect as being very bright. It may be that neither one of these "socially" higher occupations actually yields more than $3000 a year, but both of them lead to a wider view point of the possibilities of making money than is the case in a laborer's home. The salesman sees men who make large earnings in sales occupations which he considers practically the same as his own; and the merchant running a small business of his own looks forward to the career of a Wanamaker as his ultimate goal. Their ideas are passed along to their children.

Neither of these explanations covers fully the reasons for this apparent relationship between occupations of fathers and permanence of service on clerical jobs. Such factors are involved as the need for contributing to the support of the family, financial resources upon which to rely, and the less adequate education of children from the homes of unskilled workers. We feel certain, however, from our experience with clerical groups, that social status and brightness of prospect play a prominent rôle in the problems of turnover.

NOTES AND NEWS

The National Committee of the Eighth International Congress of Psychology, September 6 to 11, 1926, has announced that the following have consented to introduce symposia on the subjects mentioned before their names. The definite decisions of some others, who have accepted provisionally, will soon be published.

Understanding and Explanations: Binswanger, Kreuzlingen; Jaspers, Heidelberg; Spranger, Berlin.

Intensity of Sensations: Boring, Cambridge, Mass.; Myers, London; Werner, Hamburg.

Shape-qualities: Benussi, Turin; Koffka, Giessen; Michotte, Louvain.

Behaviorism: McDougall, Cambridge, Mass.; Pieron, Le Vésinet.

Psychology of the Primitive Races: Bartlett, Cambridge; Levy-Bruhl, Paris; Mayer-Gross, Heidelberg; Storch, Tübingen.

Religious Psychology: Janet, Paris; Leuba, Bryn Mawr.

Beside these the following ladies and gentlemen have expressed their intention to lecture on some special subject:

Miss Baumgarten, Solothurn; Mrs. Bühler, Vienna and Messrs. Bühler, Vienna; Buyse, Louvain; Foucault, Montpellier; Jaensch, Marburg; Juhász, Budapest; Ley, Brussels; Marbe, Würzburg; Mayer-Gross, Heidelberg; Paulhan, Paris; Peters, Jena; Rignano, Milan; Rupp, Berlin; Selz, Mannheim; Spearman, London; Stern, Hamburg; Thurstone, New York; Werner, Hamburg.

Those desiring to attend the Congress are requested to send names and addresses with subscription (25 sh. per member) before January 1, 1926, to Dr. F. Roels, Maliebaan, 86, Utrecht, Holland.

At the meeting of the Taylor Society, December 2 to 5, 1925, in New York City, the following papers were offered: "The Present State of Industrial Psychology," by Lillian M. Gilbreth; "Psychology in the Organization of Prison Industries," by Edgar A. Doll; and "Purpose as a Psychological Factor in Management," by Ordway Tead. Each paper was discussed by leading psychological, prison and business experts.

The Eye Sight Conservation Council of America, Times Building, New York City, reports that defective eyesight is affecting the country's industrial output, is handicapping education and is a growing menace to human welfare. The deleterious influence of eye defects and eye diseases is a challange to civilized effort in social control. The results of original research show that defective vision is wide-spread among industrial workers and school children and a prolific source of waste in both industry and education. Those interested in bringing about a reform in this matter are requested to communicate with this Council at the address given above.

The Winnetka Graded Book List, to be published in December by the American Library Association, Chicago, Ill., is a scientific study of children's tastes in reading which includes a list of 700 books which boys and girls have read and enjoyed, annotated with the children's own comments and based on the opinions of 36,750 children in thirty-four cities and towns located from Quincy, Mass., to Santa Fe, New Mexico. All types of schools were represented as well as a wide range of nationalities. Such a list should be of material aid to parents, librarians and teachers in selecting suitable and interesting books for each individual child.

"Basic Considerations in the Organization of Special Education for Mentally Deficient and Retarded School Children" is the title of the first report formulated by The Ohio Institute, assisted by the Central Committee on Special Education of Ohio, under the chairmanship of Miss Hannah L. Protzman, of The Ohio Institute, Columbus, Ohio.

Among the topics given consideration are the following:

1. Growing demand for the organization of special classes for mentally deficient children in the public schools.
2. Lack of coördination of special education in Ohio.
3. Organization of state committees on special education and programs of work.
4. Basic considerations to be recognized in the organization of special education for mentally deficient and retarded school children.
5. Organization of special classes in school districts having no psychological facilities.
6. Advisory service afforded through The Ohio Institute.

The Massachusetts Society for Mental Hygiene announces the first number of a new bulletin, "The Human Factor," which is expressly prepared for the busy executive in business or industry. It will appeal strongly to employers, personnel workers, educators in stores or factories, management executives, employment managers, physicians and all others who are in any way concerned with the human factor in business or industry. "The Human Factor" is under the editorial direction of Dr. Henry B. Elkind, Medical Director of the Massachusetts Society for Mental Hygiene and an unusually able consulting board. It will be sent free to members of the Massachusetts Society for Mental Hygiene and on request to all residents of Massachusetts who are interested.

BOOK REVIEWS

OSKAR PFISTER. *Love in Children and Its Aberrations.* Dodd, Mead and Company, New York, 1924. Pp. 576. Price $7.50.

This volume is translated from the German by Eden and Cedar Paul. It is the stated purpose of its author to make known to the general reader "the assured and fruitful results of the modern study of unconscious mentation, and of the unconscious in mental development." However, he assures the reader that it is no purpose of the book to supply an introduction to the practice of psychoanalysis. He insists that he hopes to lead teachers and parents to a correct understanding of the development of love in children to the end that its desirable needs may be cared for. He further believes that the awakening of love will be the solution for the ills of modern life.

In chapter one the author gives a brief survey of the history of the problem of love. He presents the consideration of thinkers from Akhenaton writing in the fourteenth century B.C. to Duboc in the twentieth century A.D. In these more than three thousand years of effort to understand love, no final and satisfactory solution is to be found. In this connection the author quotes Ibsen as saying: "Never before was a word so full of falsehood and fraud as the little word 'love' has become to-day." In the past some have found love to be an aspect of the intellect; others have found its origin in the emotions; while still a third group have made it a matter of the will. On the question of the origin and nature of love, the author gives us the following: "Love is a feeling of attraction and a sense of self-surrender, arising out of a need, and directed towards an object that offers hope of gratification."

In a brief chapter the author gives a sketch of the treatment of love in children by the great educators of history. The main body of the volume is divided into three parts each discussing a general topic. In this portion of the book we find the following questions considered: The love of the child for its parents; the love of the child for other persons than his parents; the love of the child for persons and things outside the home; the love of animals, nature, the acquirements of civilization and social institutions, the love of the child for himself and for God and divine things. There is also a consideration of the formative forces and experiences of love as well as the training of love in children and the treatment of its disorders.

The volume concludes with a fair bibliography and index. In closing, the reader is reminded that "the love sentiment and education to love must constitute the kernel of all our educational work." It must be saved from imprisonment in the unconscious. This can in a large measure be realized through proper and adequate diagnosis and the application of the best in mental hygiene. In short, we have in this volume an endeavor to apply the method of Freudian psychology to a consideration of one of the most subtle elements of human life. How well the author succeeds depends very much upon the reader's conception of the unconscious and the possibility of making its nature and functions known.

WILLIS L. GARD,
Ohio University.

L. L. BERNARD. *Instinct—A Study in Social Psychology.* Henry Holt and Company, New York, 1924, 550 pp.

This book is a compendium of the author's findings upon the problems of instinct after fifteen years of critical and constructive inquiry. His essential thesis is that instincts do not dominate the formation of habit, as is usually supposed, and that instincts are not the chief determinants in molding the character of social institutions. Against the instinctivists Professor Bernard contends that the environmental influences, particularly the socio-psychological, are the most potent; and in this he is in full agreement with John Dewey, F. H. Allport, C. C. Josey and other recent writers upon the subject. The treatment is in the main negative, most of the space being devoted to exposing the false and conflicting uses of the term instinct. Associated with these confusing uses of the term instinct, the author argues, are the erroneous psychological theories of emotions and sentiments. Professor McDougall's position is typical of those theories brought under criticism.

The assumption upon which Professor Bernard's study rests is that "instincts and habits are defined in terms of their neural structure rather than of their value to the organism or society, and that they are to be distinguished from each other by determining their respective origins" (p. 5). The approach to the problem of the relative importance of instinct and habit for the individual and society, he holds, is through embryology, behavioristic psychology, and sociology. Instinct is defined as an inherited action pattern, as an activity process based on a neural mechanism (p. 150), as "an inborn (in the sense of inherited) process which has remained intact, that is, which has not been remade through the process of learning or of making new adjustments by means of the substitution of new stimuli or responses for the old which were inherited" (p. 84). On what is inherited he says: "Any characteristics received after the process of fertilization has taken place, that is, not contained in the chromosomes of the uniting parent cells, is not inherited. It is acquired" (p. 259).

Further illuminating portions of the study are:

1. The discussion of the manner in which such subsidiary adjustment mechanisms as reflective thought, speech and language develop as phases of habit. Language becomes externalized through writing and thus as potential habit or thought stored outside the organism, it becomes an effective environmental control process. Social institutions, also products of experience, assist in the domination of human adjustment processes, leaving instinct to play a secondary rôle. 2. The tabulation of 5084 specific instincts, a classification of these as given by standard authors, books, and cases,—all of which clearly and conclusively proves the absence of critical standards in the use of the term. 3. The analysis of what many writers call instincts of complex character, as, e.g., maternal, paternal, play, fighting, construction and acquisition. All of these and others are shown to be habit complexes with only a remote instinctive basis. 4. The critical and comparative treatment of current psychological explanations of feeling, emotion, and sentiment, showing that they are not so exclusively instinctive as supposed, but rather that the acquired elements are dominant and directive.

Students of psychology and the social sciences will feel greatly indebted to Professor Bernard for this clear and cogent argument in favor of definiteness and uniformity in the use of the term instinct. Exactness and consistency in the use of the fundamental working concepts of the social sciences is quite as important as it is in the physical sciences. The author's motive throughout the argument, namely, a concern for accuracy, truth, and utility, is certainly sound. However, it is just here that a question is likely to be raised. Few, if any, will be prepared to show that he is wrong in contending that much which psychologists and social scientists call instinct is in reality habit and therefore acquired. But it will be asked, is it correct or even most useful to interpret instinct in purely mechanistic and quantitative terms? The author charges the instinctivists with being metaphysical and mystical. May they not reply with a *tu quoque?* Only in the author's case the implication is mechanistic instead of teleological. Is it established that biology and psychology can adequately interpret their respective data exclusively in mechanistic and quantitative terms? The reviewer is in doubt about the adequacy of an exclusively mechanistic psychology as a basis for the social sciences. With this rather unnecessary implication excepted, the conclusions offered are clear gain.

MAURICE T. PRICE, *Christian Missions and Oriental Civilizations.* Published privately but obtainable from the Christian Century Press, Chicago, 1924. 578 pp.

Here in the categories of social psychology is offered to the public the first thorough scientific study of the impact of Protestant Christian

Missions upon non-Christian peoples. The point of view is unique. The missionary enterprise is regarded as a natural social phenomenon to be described in a disinterested way. The source materials gathered through twelve years of observation and investigation in America and China include: autobiographies, biographies, statements and observations by native and foreign leaders and workers—Christian, non-Christian, and anti-Christian; reports and observations of travellers and ethnologists; Western criticism of Missions and apologetics for Missions; magazines and newspapers. A vast amount of thoroughly representative data is carefully classified and interpreted as a study in "culture-contact and cross fertilization of cultures" with an emphasis upon behavior aspects rather than upon the material effects of such contacts. It is in reality a critical study of the merging of Eastern and Western civilizations. Quotations from the sources are abundant and well selected. The author seeks to make his treatment objective and has chosen for this purpose the working concepts of behavioristic and experimental psychology. Convenient summaries, interpretative in character, appear at the end of many chapters.

It is impossible in a brief review to do justice to so ambitious an undertaking, and, moreover, to one so well done. The study should be of value to leaders in the Christian Church and to any others who wish to view without prejudice the diverse reactions of individuals and groups to Christian Missions. The more minute and subtle reactions as well as the more conspicuous are interestingly depicted. Foreign trade and foreign politics are usually regarded as the chief means of culture association and culture blending. This work shows that Christian missionary endeavor is an agency of equal importance.

The author labors under no delusion as to the finality of his study. He recognizes that is is but the beginning of a type of work which for more accuracy and completeness must await the appearance of more and better materials, as well as new modes of analysis and organization. The highly technical character of the language makes of it practically a closed book to any except the trained psychologist and sociologist.

JAMES MICKEL WILLIAMS. *Our Rural Heritage.* Alfred A. Knopf, New York, 1925. 246 pp.

Our Rural Heritage is a pioneer, historical treatise on the social psychology of rural life. It is a "cross sectional description of the attitudes and beliefs" of a few typical New York communities (chiefly one) from the close of the Revolution to the end of 1874. The author believes that the historical, social psychology of a typical rural community in Iowa, Illinois, or Ohio would not be essentially different. A later book, soon to be published, will treat the rural psychology of this New York community for the remaining periods— 1875 to 1900, and 1900 to the present time.

With painstaking care Dr. Williams examined the sources, as, e.g., records of town meetings, assessment rolls, census records, school and church records, the files of newspapers, and, in addition, held hundreds of interviews with old residents. The method of treatment is thus inductive, so far as any method may be inductive. Attitudes, beliefs, and rural social processes are first observed, then interpretations are given and inference drawn. The psychological effects of the physical environment, as, e.g., climate, weather, topography, fertility of the soil, accessibility to markets and social centers, are regarded as causal in the formation of such rural attitudes as self-reliance, independence, perseverance, acquisitiveness, and resignation. Attitudes appearing in the family, attitudes to neighbors, kinship, religion, law, politics, education, and government,—all are treated critically yet with sympathetic insight and understanding. Much is made of the economic motive in the formation of these attitudes and beliefs.

The closing chapter on *Our Rural Heritage and the National Life* is significant. Here the author frankly sets up the hypothesis that the fundamental attitudes and beliefs formed in the centuries of our dominant rural life have been carried over and up into our more recent industrial and professional mode of life with little or no criticism. In this way what may have been a virtue under one mode of life becomes a vice in another and different mode of life. The author believes that before we can understand the modern agrarian movement, and, before a restoration of economic and political stability in American life can take place, serious effort must be made to understand and appreciate the basic psychological processes of rural life.

As in the case of all pioneer studies, this one will no doubt be subjected to severe criticism. However, the reviewer believes that it is all that it purposes to be,—an honest, scholarly attempt not merely to enumerate rural traits or to make a social survey, but rather to interpret psychologically and genetically rural attitudes and beliefs in respect to their causes and effects. It should be in the hands of all students of rural problems.

JOHN DEWEY. *Experience and Nature.* Open Court Publishing Company, Chicago, 1925. 443 pp. $3.00.

This book comprises ten lectures, the first series, given upon the Paul Carus Foundation. It was especially fitting that Professor Dewey should be chosen by the American Philosophical Association to deliver these lectures, since he is the leader of that school of philosophic thought most uniquely American. The author here applies his empirico-psychological method to the task of interpreting the deeper meaning of Experience, Nature, Existence, Life, Mind, Matter, Ends, Art, and Value. The point of view is genetic, behavioristic and functional.

With a positive distaste for cleavages, splits and dualisms, e.g., "between sensuous appetite and rational thought, between the particular and the universal, between the mechanical and the telic, between experience and science, between matter and mind," between means and ends, the universe with man and his achievements is viewed as of one piece, one stream, one process. Our attention is focused now upon this phase of the process, now upon that. Things *are* what they do; they are what their histories reveal them to be. Meanings and values are instrumental always, whether in science or in art. Thus the working concepts are essentially those of biology and psychology. This approach, the author believes, is concrete and scientific, the only adequate means of resolving age-long difficulties and the conflicting theories of warring schools.

WALTER S. GAMERTSFELDER,
Ohio University.

DOUGLAS FRYER. *Vocational Self-Guidance.* J. B. Lippincott Co., 1925. 385 pp.

Dr. Fryer presents this book as a result of practical experience in the vocational office. He offers it as a plan of vocational self-guidance for the intelligent man or woman. The book is not intended for the guidance of young people. The occupational information in Part II can, however, be used to advantage with the young.

The author emphasizes a fact which has often been previously stated, viz., that the individual must make his own choice of occupation. He urges the individual to investigate vocational opportunities and to study his own personality.

Dr. Harry Kitson in his introduction to the book says that "while trying to make the procedure of vocational guidance as nearly scientific as possible we must nevertheless recognize that at present it is only in a common-sense stage of development. Doctor Fryer recognizes this. He was obliged to do so while serving as a practical vocational counselor. His plans proposed in this book have, therefore, the merit of being sensible and practicable. At the same time it is his aim as a psychologist to give the work as much scientific foundation as possible. And this goal has not been lost sight of in the book. Indeed, vocational guidance is a field that requires supremely a combination of the common-sense and scientific points of view."

In spite of the author's efforts to keep the book on the best possible scientific basis, he seems to have neglected to warn the reader that the individual's judgments of his own traits and abilities are not very reliable and that many serious errors will probably accompany any plan of self-guidance. This would be particularly true of the less intelligent reader. The author should recognize more emphatically that the judgments and assistance of others should be sought by the one who is attempting self-guidance.

Part II of the book is composed of Chapters 6 to 27 which form a series of very excellent and informing discussions of various occupations. These are contributed by persons who are experts in their fields. For the counselor and teacher this is perhaps the most usable part of the book.

Part III on "The Business Professions for Women" is very brief and written very distinctly from the point of view of the woman's right and need to engage in various types of occupational life. The author states that woman "will have many obstacles and difficulties to overcome, perhaps greater and certainly different, from those confronting the man of equal ability, but all the possibilities of success are hers and all the rewards, of money, of power, of self-expression, of service, if she but wants them, not merely enough to dream for them or to talk of them, but to *work* for them, to plan and prepare for them, and to believe both in them and in herself."

Part IV is entitled "Getting the Job." It deals helpfully with the selecting of the job and sales methods.

The book as a whole contains material which can be used particularly in the guidance of adults. It could be put into the hands of the adult himself, although as the author says, it would have to be an intelligent adult who could use it to advantage. His best use of it, however, would be in coöperation with the expert counselor or psychologist.

THOMAS C. MCCRACKEN,
Ohio University.

L. L. THURSTONE. *The Fundamentals of Statistics.* New York, The Macmillan Company, 1925. 237 pp.

ARTHUR S. OTIS. *Statistical Method in Educational Measurement.* Yonkers-on-Hudson, New York, World Book Company, 1925. 337 pp.

HAROLD RUGG. *A Primer of Graphics and Statistics for Teachers.* Boston, Houghton Mifflin Company, 1925. 142 pp.

Early books on statistical method were written by scholarly scientists, usually well-trained in mathematics, and were read and understood by a limited number of individuals whose interest in statistics and whose knowledge of mathematics permitted them to follow the development of the subject. With the extension of statistical methods from the original field of biology to the fields of economics, psychology, sociology, and educations, however, it soon became apparent that statistical treatises, presupposing no more mathematical training than is given in the secondary school, must be prepared. That the efforts of Thorndike, the Eldertons, King, Rugg, Secrist, Jones, and others, to meet this need were only partially successful is witnessed by the large number of statistical books published in the last two or three years. In addition to the three under review, we have Odell, Mills, Kent, Riegel, Forsyth, Jerome, and others. The outstanding scholarly work of the present day—Kelly—and the technical handbook by Rietz, *et al.* are standard reference works for the more advanced student.

Of the three books under consideration, Otis and Rugg were written primarily for the profession of education while Thurstone is a more general treatment, although the editor—O'Shea—in his introduction assumes that the book will be read chiefly by teachers. Otis and Thurstone are considerably more elaborate discussions than is Rugg, probably occupying, so far as reading difficulty is concerned, a position about midway between Rugg's two books—his *Statistical Methods Applied to Education* and his *Primer*. Each of the three books naturally possesses some unique features.

Rugg is brief, concise, and reasonably clear. The usual measures of central tendency and variability are briefly described and well illustrated. Simple correlation is computed by three methods. The book is profusely illustrated; it contains 115 figures in its 142 pages. (Thurstone has 40 figures in 237 pages and Otis, 68 figures in 337 pages.) For teachers' meetings and short course normal school classes, this book will be popular and justly so.

Thurstone is brief and non-technical but makes much better provision for an understanding of fundamental mathematical concepts than is usual. The book has chapters on Linear Relations, Non-linear Relations, The Equation of a Straight Line through the Origin, The General Equation of a Straight Line, The Binomial Expansion, and The Probability Curve, as well as the usual discussion of the tabulation of data, graphic representation, and measures of central tendency and position, variability, and relationship. The discussions are adequate for an elementary course but clear and comprehensible. Unfortunately, the book is not supplied with references or a bibliography and makes no mention of partial or multiple correlation.

Otis has prepared the most elaborate book of the three. Explanations are more detailed in character, and probably more lucid to some beginners, than is the case in either Rugg or Thurstone. In the reviewer's judgment, the discussion and interpretation of correlation is not only the best feature of the book but the best available. Interpretation is accomplished by numerous examples and illustrations and by introducing Kelly's *coefficient of alienation*, a concept of great importance and value. Partial and multiple correlation are briefly treated. In the opinion of the reviewer, the book is overloaded with discussions of percentile curves and graphs and is faulty in that illustrative test material comes largely from the publications of one company.

With the appearance of these three books it should no longer be necessary for educators to be unacquainted with elementary statistical analyses and interpretations.

R. L. Morton,
Ohio University.

Stanley High. *Europe Turns the Corner.* New York, The Abingdon Press, 1925. Pp. 308. $2.00.

The author bases his work on observation and first hand contacts with the changing and improving Europe. As a student of political and economic conditions during several trips to Europe in the last five years, he concludes that the method and spirit of war have given place to peaceful overtures and that the program of revenge has been superceded by a policy of co-operation. Although the issues have not been settled, they have ceased to be threats of war.

The British labor government under Ramsay MacDonald was the beginning of the "turn;" the program of Mussolini in Italy, although unfortunate in its international aspects, has contributed through its worthy domestic features; the French have abandoned their policy of revenge and their ambition to destroy Germany; and the new Russia is emerging through education, religion, and intercourse with other peoples. The author sees in the Baltic League, the Little Entente, the financial and economic co-operation with Hungary, and other movements the development of attitudes which may make possible a United States of Central Europe.

Without the United States, Europe proceeds in the League. Indications of a transformed Europe are noted in the labor government in England, the downfall of Poincare in France, the London Conference, the Dawes plan, the recognition of Russia by England and France, and the pronouncement of the Protocol of the League. Besides an account of the development of these factors, the author presents an analysis of them and their implications in international affairs. The significance of these interests for the security and prosperity of the United States is stressed.

Events of 1925, such as the Locarno Conference with its hopeful prospects, confirm the author's optimistic position that European conditions offer the bases for a new structure of world society. It remains to be observed whether the concept of the world as a social group can be retained for a period long enough to enable such attitudes to become habitually established. Colonel Edward M. House, who has written the introduction, concludes: "He has written his story well and understandingly, and coming generations who would know of these times will be his grateful debtors."

Psychology and History. HARRY ELMER BARNES, Professor of Sociology in Smith College. The Century Co., New York. 1925. Pp. 195.

The author brings to this study the rich experience of the historian and the sociologist who has been influenced also by stimulating contact with the late President G. Stanley Hall of Clark University. The dedication of his more extensive work "The New History and the Social Studies," of which this book is the reprint of two chapters, indicates the author's debt to the "Columbia School of Historians" which was led by James Harvey Robinson. The book presents "a serviceable summary of the major developments in social psychology which are significant for the historian," and for the psychologist "a review of the efforts which have been made by historians . . . . to apply psychological methods and data to the interpretation of the history of man and human culture." The author gives sarcastic evidence of irritation with "respectable" historians' viewpoint that their only service is to make the past known without concern whether that past has any serviceable connection with the present. He believes that the interpretation of history is dependent upon an understanding of the principles of individual and of social psychology. Although the former is valuable in interpreting the personalities of historic periods, the latter is necessary to understand the development of the social group. The author concludes that "at the same time that social psychology was being elaborated there was developing a genetic view of history, the idea of continuity of history, which is based upon the knowledge that our present cultural complex is primarily the product of a long inheritance from past conditions and that the present can be understood only when viewed in the light of its historical antecedents. When this attitude was accepted by the more progressive students . . . . social psychology had put at their disposal the methodology and point of attack which was necessary to show how the modern age had come about . . . . if the collective psychology was an all-important factor in modern life it should have had a great, if not equal, significance in all ages. The systematic following up of this clue has constituted essentially the psychological interpretation of history in so far as it has been developed." The reader is introduced to the psychology of instincts, behaviorism, and the psychology of the unconscious as the more pronounced innovations for the consideration of those concerned with the psychological interpretation of history.

After a brief "Prefatory Note" the author introduces, with generous quotation, the conspicuous contributors to the fields of psychology and history who have discussed the more important aspects of the development of psychology in relation to history and the social studies. The author has arranged his materials to demonstrate that the psychologist has made contributions which may be utilized by the historian to produce a utilitarian type of history. The illustrative materials, calculated to demonstrate the principles presented, admirably suggest the values of the psychological approach. To the psychologist, this study presents a summary of the work of those who have tried to utilize psychological contributions in producing historical writing; and to the historian, it lends encouragement to apply the principles of psychology to the interpretation of the past for the understanding of the present.

EDWIN B. SMITH,
Ohio University.

NEW BOOKS AND PAMPHLETS RECEIVED

Books and pamphlets for review should be sent to James P. Porter, Department of Psychology, Ohio University, Athens, Ohio.

L'Avenir du Travail. Bulletin de l'Association Internationale pour la Protection légale des Travailleurs. Vol. III, Nos. 1-2. 1925. Imprimerie, Librarie Berger-Levrault, Paris. 75 pp.

Contributions to Education, Teachers College, Columbia University:

- *A Comparison of the Intellectual and Educational Status of Neurotic and Normal Children in Public Schools.* Esther Katz Rosen. No. 188.
- *Comprehension Difficulties of Ninth Grade Students in the Study of Literature.* Theo. W. H. Irion. No. 189, 1925.
- *An Experimental Method for the Discovery and Development of Tests of Character.* Theodore P. Lentz, Jr. No. 180, 1925.
- *The Measurement of Fair-Mindedness.* Goodwin B. Watson. No. 176, 1925.
- *The Opposites Test.* Andrew Tennant Wylie. No. 170, 1925.
- *A Study of the Relationship Between Rate and Ability.* Lillian May Hunsicker. No. 185, 1925.

The Constructive Ability of Young Children. Lovisa C. Wagoner. University of Iowa Studies, First Series No. 94, June, 1925. Iowa City. 55 pp.

The Emotions of Young Children. Leslie R. Marston. University of Iowa Studies, First Series No. 95, June 15, 1925. Iowa City, Iowa. 99 pp.

Europe Turns the Corner. Stanley High. The Abingdon Press, New York. Price $2.00. 308 pp.

Health and Safety in the New Curriculum. Geo. E. Payne and Louis C. Schroeder. The American Viewpoint Society, Inc., New York, 1925.

An Introduction to Objective Psychology. G. V. Hamilton. The C. V. Mosby Company, St. Louis, Mo., 1925. Price $5.00.

Mental Measurement in Educational and Vocational Guidance. Harvard Bulletins in Education. Cambridge, Mass. No. X, November, 1924.

Modern Theories of the Unconscious. W. L. Northridge. E. P. Dutton & Co., New York City. Price $3.75. 192 pp.

Occupational Efficiency of the Mentally Defective. Bulletin of the University of Minnesota, College of Education. Vol. 27, No. 55, December 20, 1925.

Preliminary Report on Supervision in County Demonstrations. Educational Bulletin No. 74. Dept. of Public Instruction, State of Indiana, 1924.

A Primer of Psychology. Edward Bradford Titchener. The Macmillan Co., New York. 1925.

Psychology and History. Harry Elmer Barnes. The Century Co. New York. Price $1.50. 195 pp.

The Recent Foreign Policy of the United States. George H. Blakeslee. The Abingdon Press, New York. Price $2.00. 368 pp.

Religious Education Survey Schedules. Walter S. Athearn. Vol. III. George H. Doran Co., 1924.

Reorganization and Administration of an Elementary School to Meet the Needs of a Community. University of California Publications in Education. Vol. 2, No. 2, Berkeley, Calif., 1925.

Report on the Present Position of Vocational Guidance and Vocational Selection. M. Boole Stott. The Women's Employment Publishing Co. Ltd., London, England.

Report of the Survey of the Public Schools of Philadelphia. The Pennsylvania State Department of Public Instruction, Four Books, 1922.

Special Disabilities in Learning to Read and Write. Elizabeth E. Lord, Leonard Carmichael and Walter F. Dearborn. Harvard Monographs in Education, Series I, Vol. II, No. 1. The Graduate School of Education, Harvard University, Cambridge, Mass. Price $1.00. 76 pp.

Studies of Mental Defects and Handicaps. J. E. Wallace Wallin. Miami University Bulletin, Series XXII, No. 5. Oxford, Ohio. Price 75 cents. 177 pp.

A Study of Supervised Study. William Arthur Brownell. University of Illinois Bulletin, Vol. XXII, No. 41. Urbana, Ill. Price 50 cents. 48 pp.

The Taxation Amendment. J. W. Tannehill. Superintendent of the Building and Loan Department, Room 409, State Office Building, Columbus, Ohio. 10 pp.

Tryout Set, McQuarrie Test for Mechanical Ability. Associated Students Store, University of Southern California, Los Angeles, Calif.

The Visiting Teacher in Rochester Report of a Study. Mabel Brown Ellis. Joint Committee on Methods of Preventing Delinquency, 1925.

Department of the Interior Publications, Department of the Interior, Washington, D. C.:

- *Accredited Secondary Schools in the United States.* Frank M. Phillips. Bulletin No. 11, 1925. Price 15 cents. 110 pp.

Helps for the Rural School Nurse. HARRIET AND HAZEL WEDGWOOD. Health Education Bulletin No. 17, 1924.

Introduction of Algebra into American Schools in the Eighteenth Century. LAO GENEVRA SIMONS. Bureau of Education Bulletin, No. 18, 1924.

Land-Grant College Education, 1910–1920. Part III. WALTON C. JOHN. Bulletin No. 4, 1925. Price 25 cents 108 pp.

List of Publications Available September, 1925. 24 pp.

List of References on Higher Education. Bureau of Education, Library Leaflet, No. 28, December, 1924.

Professional Staff of State Departments of Education. ARTHUR WESLEY FERGUSON. Bulletin No. 17, 1925. Price 10 cents. 64 pp.

The School as the People's Clubhouse. HAROLD O. BERG. Physical Education Series No. 6, 1925.

Statistics of Kindergartens, 1923–24. FRANK M. PHILLIPS. Bulletin No. 20, 1925. Price 5 cents. 7 pp.

Statistics of State Universities and State Colleges, for Year Ending June 30, 1924. Bulletin No. 12, 1925. Price 5 cents. 23 pp.

Teacher and Pupils' Reading Circles. ELLEN C. LOMBARD. Home Education Circular No. 7, March, 1925.

INDEX OF AUTHORS

The names of authors of original contributions are printed in CAPITALS AND SMALL CAPITALS.

ALLEN, IRVING, 367.
ARLITT, ADA HART, 215.
ARTHUR, GRACE, 300.
BATES, ROBERT L., 95.
BILLS, MARION A., 424.
BRANDENBURG, GEORGE C., 130, 281.
BURTT, H. E., 5.
CLEETON, GLEN U., 101.
DOBELL, E. M., 5.
DOWD, CONSTANCE E., 215.
FREYD, MAX, 349.
FREYER, DOUGLAS, 304.
Gamertsfelder, Walter S., 99, 335, 435.
Gard, Willis L., 323, 431.
GARTH, THOMAS R., 382.
Gentry, J. R., 326.
Good, H. G., 333.
GUILFORD, J. P., 196.
HOITSMA, RALPH K., 293.
HOLLINGWORTH, H. L., 122.
HOLMES, JOSEPH L., 60.
HULL, CLARK L., 327.
HYDE, WINIFRED F., 196.
Jefferson, B. L., 324.
Kantor, J. R., 94.
KITSON, HARRY D., 104, 209, 329, 367.
Laird, Donald A., 96, 206.
LANGLIE, F. A., 339.
LOWE, GLADYS M., 53.
McCracken, Thomas C., 436.
MANSON, GRACE E., 156, 230.
MAY, MARK A., 29.
MINER, J. B., 357.
Morton, R. L., 438.
MURDOCH, KATHARINE, 364.
NIXON, H. K., 176.
OTIS, FRANCES E., 311.
PATERSON, DONALD G., 339.
POFFENBERGER, A. T., 22.
Pratt, C. C., 211.
REMMERS, H. H., 206, 417.
RORISON, W. A., 225.
Rowles, Emmett, 94.
SHELLOW, SADIE MYERS, 69.
SHEN, EUGENE, 66.
SHIMBERG, MYRA E., 53.
Smith, E. B., 440.
SOLINGER, S. B., 225.
STINCHFIELD, SARA M., 109.
STRONG, EDWARD K., JR., 75.
SUNNE, DAGNY, 256.
SYMONDS, PERCIVAL M., 188.
THOMPSON, L. A., JR., 417.
THORNDIKE, E. L., 1.
VITELES, MORRIS S., 131.
WARDEN, C. J., 225.
Warner, M. Lavinia, 92.
YOUNG, HERMAN H., 371.

INDEX OF SUBJECTS

Titles of books and pamphlets are followed by the name of the author of the work reviewed. All other references are to original articles.

Achievement and success in life, school, 22; —— personality and vocational, 281.
Activity conditioned by emotional states, motor, 417.
Adolescents, personality tests, white and negro, 256.
Advertisements, attention to, 176.
Advertising, free association method a measure of the efficiency of, 60; —— pictures of people in magazine, 367; —— tendencies in use of headlines in magazine, 225; —— material, curve of forgetting for, 5.
Alternative forms of examinations equal in difficulty, provision of, 1.
An Annotated Bibliography Dealing with the Classification and Instruction of Pupils to Provide for Individual Differences (Odell), 320.
Association method a measure of efficiency of advertising, free, 60.
Attention to advertisements, two studies, 176.

Brain Tests (Saunders and Palmer), 200.

Cases in Retail Salesmanship (Kneeland), 207.
Chemistry, classification test for students in, 196.
Child, The: His Nature and His Needs (O'Shea), 91.
Christian Missions and Oriental Civilizations (Price), 432.
Classification of chemistry students, test for, 196.
Clerical worker and his permanence on job, social status of, 424.
Clinic case with misconduct a by-product, speech, 371.
Colgate mental hygiene test, reliability and relationship of, 293.
College women, speech of five hundred freshmen, 100.
Constant and Variable Errors of Educational Measurements (Monroe), 320.

Differences in intelligence tests, group, 156; —— in intelligence test performance, personality, 230.
Difficulty of types of questions, relative; group differences in intelligence tests, 156.
Disparity between scholarship and intelligence, prediction of, 357.
Dynamic Psychology an Introduction to Modern Theory and Practice (Moore), 101.

Effects of supervised play and formal gymnastics, relative transfer, 215.
Efficiency of advertising, free association method a measure, 60.
Emotional states, motor activity as conditioned by, 417.
Eugenics, Genetics and the Family, 92; —— in Race and State (Second International Congress of Eugenics), 92.
Europe Turns the Corner (High), 438.
Examinations equal in difficulty, provision of alternative forms, 1.
Experience and Nature (Dewey), 434.
Experiment on effect of note-taking in lecture courses, suggestion for, 69.

Fables as moral judgment test, critique of, 53.
Forgetting for advertising material, curve of, 5.
Formal Gymnastics, relative transfer effects of supervised play and, 215.
Freshmen college women, speech of five hundred, 109.
Full-blood Indians, intelligence of, 382.
Fundamentals of Social Psychology (Bogardus), 98.
Fundamentals of Statistics (Thurstone), 436.

Gymnastics, relative transfer effects of supervised play and formal, 215.

Headlines in magazine advertising, use of, 225.
Helmholtz's Treatise on Physiological Optics Vol. I and II (Southall), 209, 210.
Hugo Münsterberg: His Life and Work (Münsterberg), 332.
Hygiene test, reliability and relationships of Colgate mental, 293.

Improvement of the Written Examination (Ruch), 99; 205.
Indians, intelligence of full-blood, 382.
Influence of Tobacco Smoking on Mental and Motor Efficiency (Hull), 95.
Instinct—A study in Social Psychology (Bernard), 431.
Intelligence, prediction of disparity between scholarship and, 357; —— test performance, personality differences in, 230; —— tests, group differences in; the relative difficulty of types of questions, 156.
Interest, note on vocational, 304.
Intoxicated, when is a man, 122.

Job, social status of clerical worker and his permanence on, 424.
Judgment test, critique of fables as moral, 53.

Lecture courses, suggestion for experiment on effect of note-taking in, 69.
Leopold and Loeb (Urstein), 95.
Life and Confessions of a Psychologist (Hall), 330.
Love in Children and Its Aberrations (Pfister), 430.

Magazine advertising, pictures of people in, 367; —— advertising,

tendencies in use of headlines in, 225.
Measure of efficiency of advertising, free association, 60.
Measuring the Results in Education (Trabue), 324.
Memory in its bearing upon mental tests, phenomenal, 311.
Mental tests, phenomenal memory in its bearing upon, 311; —— Hygiene test, reliability and relationships of Colgate, 293.
Mind as Behavior (Singer), 96.
Misconduct as by-product, speech clinic case, 371.
Moral judgment test, critique of fables as, 53.
More Abundant Life, or Bio-Psycho-Genetics (Taylor), 96.
Motor activity conditioned by emotional states, 417.

Negro adolescents, personality tests: white and, 256.
Note-taking in lecture courses, experiment on effect of, 69.

Our Rural Heritage (Williams), 433.

Passing of the Phantoms, The (Patten), 325.
Performance scale, new point, 390; ——, personality differences in intelligence test, 230; —— tests, comparative value of nine, 364.
Permanence on job, social status of clerical worker and his, 424.
Personal ratings, influence of friendship upon, 66.
Personality, analyzing, 139.
Philosophy of Character, The (Pierce), 97.
Play and formal gymnastics, relative transfer effects of supervised, 215.
Point performance scale, new, 390.
A Primer of Graphics and Statistics for Teachers (Rugg), 436.
Primitive Mentality (Levy-Bruhl), 94.
Primitive Religion (Lowie), 334.
Principles of Education (Chapman and Counts), 322.
Psychology in Business Relations (Snow), 327.
Psychology and History (Barnes), 439.

Questions, relative difficulty of types of intelligence test, 156.

Rating, notes on, 188.
Ratings, influence of friendship upon personal, 66.
Relationships of Colgate mental hygiene test, reliability, 293.

Scale, new point performance, 390.
Scholarship and intelligence, prediction of disparity between, 357.
Scoring of true-false tests, empirical data on, 339.
Selection, clinical viewpoint in vocational, 131; ——, statistical viewpoint in vocational, 349.
Selling, theories of, 75.
Sigmund Freud: His Personality, His Teaching and His School (Wittels), 323.
Statistical Method (Kelley), 632.
Statistical Method in Educational Measurement (Otis) 436.

Status of the clerical worker and his permanence on job, 424.

Students in chemistry, test for classification, 196.

Studies of attention to advertisements, two, 176.

Success in life, school achievement and, 22.

Supervised play and formal gymnastics, transfer effects, 215.

Survey of the City Schools of Marion, Ill. (Monroe), 320.

Temperament tests, present status of will-, 29.

Test, critique of fables as moral judgment, 53; —— performance, personality differences in intelligence, 230; ——, reliability and relationships of Colgate mental hygiene, 293.

Tests, study of comparative value of nine performance, 361; ——, empirical data on scoring of true-false, 339; 339, group differences in intelligence; relative difficulty of types of questions, 150; ——, phenomenal memory in its bearing upon mental, 311; ——, present status of will-temperament, 29; ——: white and negro adolescents, personality, 256.

Training in the Technique of Study (Monroe and Mohlman), 321.

Transfer effects of supervised and formal gymnastics, relative, 215.

True-false tests, empirical data on scoring, 339.

Types of questions, relative difficulty; group differences in intelligence tests, 150.

Vocational achievement, personality and, 281; —— interest, note on, 304; —— selection, clinical viewpoint in, 131; —— selection, statistical viewpoint in, 349.

Vocational Self-Guidance (Fryer), 435.

White and negro adolescents, personality tests, 256.

Will-temperament tests, present status of, 29.

Women, speech of five hundred freshmen college, 109.

Work, types of, 304.

Worker and his permanence on job, social status of clerical, 424.